

### Procédé pour le diagnostic/pronostic du neuroblastome

La présente invention concerne un procédé pour le pronostic du neuroblastome.

- 5 Le neuroblastome est en fréquence la deuxième cause de tumeur solide chez l'enfant après les tumeurs cérébrales. Le neuroblastome est le plus fréquent des cancers de l'enfant avant cinq ans et représente environ 15 % des cancers avant cet âge.

Les neuroblastomes sont des tumeurs malignes développées à partir de neuroblastes nés de la crête neurale et migrant pour former les ganglions sympathiques et la médullo-surrénale  
10 durant la période embryonnaire et fœtale.

Lorsqu'un premier examen clinique permet de suspecter un neuroblastome (boule, hématome, endroit douloureux, difficulté à bouger les membres, etc...), un bilan complet est réalisé afin de confirmer le diagnostic.

Généralement, ce bilan comprend :

- 15       ➤ des examens par prélèvements (sang, urines),  
          ➤ différents examens radiologiques qui ont pour but de bien situer la tumeur, ses limites et sa taille (scintigraphie, échographie et/ou scanner et/ou IRM ),  
          ➤ des examens au microscope de fragments de tumeurs afin de découvrir exactement de quel type de tumeur dont il s'agit.

- 20 A l'heure actuelle, il n'existe pas de traitement universel lorsqu'un neuroblastome est diagnostiqué, et un traitement spécifique doit être adapté en fonction de l'âge du patient. On distingue alors principalement les traitements loco-régionaux (chirurgie et radiothérapie) afin d'enlever ou détruire la tumeur directement à l'endroit où elle se trouve et les traitements généraux (chimiothérapie), qui agissent dans tout l'organisme du patient à la  
25 fois sur la tumeur et mais aussi là où peuvent se trouver les métastases.

Le traitement du patient peut être adapté selon le pronostic du neuroblastome et la stratégie thérapeutique peut s'avérer très différente selon le stade et la caractérisation génétique des cellules tumorales. Ainsi, dans les formes localisées dont les cellules tumorales ne portent aucun caractère de mauvais pronostic, le traitement est essentiellement chirurgical alors que

dans les formes localisées dont les cellules tumorales sont de mauvais pronostic, le traitement doit être plus agressif, reposant sur la chimiothérapie et une radiothérapie locale. Il existe à l'heure actuelle différentes classifications du neuroblastome permettant de définir de la façon la plus précise possible des groupes pronostiques. Ces groupes permettent théoriquement de définir les indications thérapeutiques de façon adaptée au risque de la maladie. On peut citer notamment la classification de l'International Neuroblastoma Staging System (Brodeur et al. (1993) J. Clin. Oncol. 11, 1466-77), qui tient compte des données anatomiques actuellement reconnues comme ayant une valeur pronostique. Selon cette classification, on distingue les stades suivants :

- stade 1 : tumeur localisée totalement enlevée macroscopiquement ; ganglions homo et controlatéraux examinés et négatifs microscopiquement
- stade 2A : tumeur unilatérale enlevée incomplètement avec des ganglions homo et controlatéraux examinés et négatifs.
- stade 2B : tumeur unilatérale avec atteinte ganglionnaire homolatérale et non controlatérale.
- stade 3 : tumeur unilatérale non opérable montrant un dépassement de la ligne médiane ou tumeur unilatérale avec atteinte ganglionnaire controlatérale ou tumeur à cheval sur la ligne médiane avec extension bilatérale par infiltration ou par adénopathie.
- stade 4 : tumeur primitive s'accompagnant d'une dissémination à distance : ganglionnaire, osseuse, médullaire, hépatique.
- stade 4S : tumeur de stade local 1 ou 2 avec une dissémination limitée au foie, à la peau ou à la moëlle osseuse. Les stades 4S sont des enfants ayant un âge inférieur à 1 an.

Actuellement, le pronostic d'un neuroblastome peut être établi par l'étude de différents facteurs:

- 1) l'amplification de l'oncogène N-myc est considérée comme un outil de référence, et est utilisée par la plupart des oncologues pédiatriques pour définir, au moment du diagnostic, les malades qui doivent recevoir une chimiothérapie intensive suivie de

greffe de moelle (Seeger et al, N Engl J Med. 1985 ; 313(18):1111-6 ; Rubie et al J Clin Oncol. 1997 Mar;15(3):1171-82.).

- 2) il existerait également une corrélation entre le pronostic du neuroblastome et le rapport des catécholamines VMA (vanilmandelic acid) / HVA (homovanillic acid),  
5 au moment du diagnostic. Dans les stades avancés, une excrétion urinaire élevée d'HVA et basse de VMA, voire normale, signerait un mauvais pronostic (Laug et al Pediatrics. 1978 ; 62(1):77-83).
- 3) l'augmentation de la ferritine sérique dans les neuroblastomes est également considérée comme un facteur de mauvais pronostic (Evans et al, Cancer. 1987 ;  
10 59(11):1853-9).
- 4) le taux de LDH (lactate deshydrogénase) pourrait également être un facteur de pronostic indépendant et prédominant pour les stades localisés I à III chez l'enfant de plus d'un an, et de façon moins importante chez l'enfant de moins d'un an avec un stade IV (Berthold et al, Am J Pediatr Hematol Oncol. 1994 ; 16(2):107-15).
- 15 Toutefois, la corrélation entre l'amplification de l'oncogène N-myc et le pronostic du neuroblastome n'est pas absolue (Maris & Matthay, J Clin Oncol, 1999, 17(7) : 2264-2279). De plus la LDH et la ferritine étant deux facteurs corrélés entre eux, la fiabilité de ces facteurs pour le pronostic du neuroblastome reste discutée (Berthold et al, 1992, Am J Pediatr Hamtol Oncol, 14(3) : 207-215). Enfin, l'utilisation du rapport VMA/HVA donne  
20 une sensibilité et une spécificité insuffisantes pour le pronostic du neuroblastome.

La présente invention se propose de résoudre l'ensemble des inconvénients de l'état de la technique en présentant un nouvel outil de pronostic du neuroblastome.

D'une manière surprenante, les inventeurs ont mis en évidence que le pronostic  
25 d'un neuroblastome peut être déterminé par l'analyse de l'expression de gènes cibles sélectionnés parmi 37 gènes tel que présenté dans le tableau 1 ci après, qui sont exprimés différenciellement selon que le patient soit de bon ou de mauvais pronostic.

**Tableau 1 - liste des 37 gènes cibles selon l'invention**

SEQ ID N°	Description de la séquence	N° Genbank
1	Flap structure-specific endonuclease 1 (FEN1)	NM_004111
2	Ubiquitin-conjugating enzyme E2C (UBE2C)	NM_007019
3	Insulin-like growth factor binding protein 7 (IGFBP7)	NM_001553
4	Collagen, type I, alpha 2 (COL1A2)	NM_000089
5	Nucleolin (NCL)	NM_005381
6	Interleukin enhancer binding factor 3, 90kDa (ILF3), transcript variant 2	NM_004516
7	cDNA FLJ30781 fis, clone FEBRA2000874, correspondant au gène codant l'adenylate cyclase 1 (brain)	AK055343 / NM_021116
8	TIF1beta zinc finger protein	X97548
9	Likely ortholog of mouse tumor differentially expressed 1, like (TDE1L)	NM_020755
10	DKFZp434D1112_s1 434 (synonym: htes3) cDNA clone DKFZp434D1112 3'	AL039831
11	v-myc myelocytomatosis viral related oncogene, neuroblastoma derived (avian)(MYCN)	NM_005378
12	Small nuclear ribonucleoprotein D2 polypeptide 16.5kDa (SNRPD2), transcript variant 1	NM_004597
13	MCM2 minichromosome maintenance deficient 2, mitotin ( <i>S. cerevisiae</i> )	NM_004526
14	RuvB-like 2 ( <i>E. coli</i> ) (RUVBL2)	NM_006666
15	Immediate early protein ETR101	NM_004907
16	RNA binding protein S1, serine-rich domain (RNPS1), transcript variant 1	NM_006711
17	Ornithine decarboxylase 1 (ODC1)	NM_002539
18	Activity-regulated cytoskeleton-associated protein (ARC)	NM_015193
19	Secretogranin II (chromogranin C) (SCG2)	NM_003469
20	Structure specific recognition protein 1 (SSRP1)	NM_003146
21	Collagen, type VI, alpha 3 (COL6A3), transcript variant 1	NM_004369
22	Small nuclear ribonucleoprotein polypeptides B and B1 (SNRPB)	NM_003091
23	Acidic (leucine-rich) nuclear phosphoprotein 32 family, member B (ANP32B)	NM_006401
24	Non-POU domain containing, octamer-binding (NONO)	NM_007363
25	Peripheral myelin protein 22 (PMP22), transcript variant 1	NM_000304
26	Small nuclear ribonucleoprotein polypeptide E (SNRPE)	NM_003094
27	KIAA0436 mRNA, partial cds	AB007896
28	Fibrillarin (FBL)	NM_001436
29	Tripartite motif-containing 2 (TRIM2)	NM_015271
30	MCM6 minichromosome maintenance deficient 6 (MIS5 homolog, <i>S. pombe</i> ) ( <i>S. cerevisiae</i> ) (MCM6)	NM_005915
31	Polypyrimidine tract binding protein 1 (PTBP1), transcript variant 1	NM_002819
32	Small nuclear ribonucleoprotein polypeptide A (SNRPA)	NM_004596
33	Creatine kinase, brain (CKB)	NM_001823
34	Erythrocyte membrane protein band 4.1-like 3 (EPB41L3)	NM_012307
35	Hypothetical protein MGC3077	NM_024051
36	Tissue alpha-L-fucosidase 1 (FUCA 1)	NM_000147



37	Secreted protein acidic and rich in cysteine (SPARC)	NM_003118
----	--	-----------

Parmi ces gènes, on peut distinguer des gènes dont la fonction est connue mais qui n'ont jamais été mis en relation avec le neuroblastome ( SEQ ID N°1 à 8 ; 12 à 16 ; 18 à 26 ; 28 ; 30 à 34 ; 36) ainsi que des gènes dont la fonction est inconnue (SEQ ID N°9 ; 10 ; 5 27 ; 29 ; 35). Il est bien entendu que si différentes isoformes de ces gènes existent, toutes les isoformes sont pertinentes pour la présente invention, et pas uniquement celles présentées dans le précédent tableau.

A cet effet, la présente invention concerne un procédé pour le pronostic du neuroblastome 10 chez un patient atteint du neuroblastome caractérisé en ce qu'il comprend les étapes suivantes :

- a. on extrait du matériel biologique d'un échantillon biologique prélevé chez le patient,
- b. on met en contact le matériel biologique avec au moins un réactif spécifique 15 choisi parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant l'une quelconque des SEQ ID N° 1 à 37, étant entendu que lorsque que le gène cible présente une séquence nucléique ayant l'une des SEQ ID N°11, 17 ou 37, on met en contact le matériel biologique avec au moins deux réactifs spécifiques choisis parmi les réactifs spécifiques des gènes 20 cibles présentant une séquence nucléique ayant l'une quelconque des SEQ ID N° 1 à 37,
- c. on détermine l'expression d'au moins un desdits gènes cibles, étant entendu que lorsque que le gène cible présente une séquence nucléique ayant l'une des SEQ ID N°11, 17 ou 37, on détermine l'expression d'au moins deux desdit 25 gènes cibles

Au sens de la présente invention, on entend par échantillon biologique, tout échantillon prélevé chez un patient, et susceptible de contenir un matériel biologique tel que défini ci après. Cet échantillon biologique peut être notamment un échantillon de sang, de sérum, de

salive, tissu, de tumeur, de moelle osseuse, de cellules circulantes du patient. On dispose de cet échantillon biologique par tout type de prélèvement connu de l'homme du métier. Selon un mode préféré de réalisation de l'invention, l'échantillon biologique prélevé chez le patient est un échantillon tissulaire, préférentiellement un échantillon de tumeur ou de moelle osseuse.

Au sens de la présente invention, on entend par matériel biologique, tout matériel permettant de détecter l'expression d'un gène cible. Le matériel biologique peut comprendre notamment des protéines, ou des acides nucléiques tels que notamment les acides desoxyribonucléiques (ADN) ou les acides ribonucléiques (ARN). L'acide nucléique peut être notamment un ARN (acide ribonucléique). Selon un mode préféré de réalisation de l'invention, le matériel biologique comprend des acides nucléiques, préférentiellement, des ARN, et encore plus préférentiellement des ARN totaux. Les ARN totaux comprennent les ARN de transfert, les ARN messagers (ARNm), tel que les ARNm transcrit du gène cible, mais également transcrit de tout autre gène et les ARN ribosomaux. Ce matériel biologique comprend du matériel spécifique d'un gène cible, tel que notamment les ARNm transcrits du gène cible ou les protéines issues de ces ARNm mais peut comprendre également du matériel non spécifique d'un gène cible, tel que notamment les ARNm transcrits d'un gène autre que le gène cible, les ARNt, les ARNr issus d'autres gènes que le gène cible.

Lors de l'étape a) du procédé selon l'invention, on extraie le matériel biologique de l'échantillon biologique par tous les protocoles d'extraction et de purification d'acides nucléiques bien connus de l'homme du métier.

A titre indicatif, l'extraction d'acides nucléique peut être réalisée par :

- une étape de lyse des cellules présentes dans l'échantillon biologique, afin de libérer les acides nucléiques contenus dans les cellules du patient. A titre d'exemple, on peut utiliser les méthodes de lyse telles que décrites dans les demandes de brevet:
  - o WO 00/05338 sur la lyse mixte magnétique et mécanique,
  - o WO 99/53304 sur la lyse électrique,
  - o WO 99/15321 sur la lyse mécanique.

L'homme du métier pourra utiliser d'autres méthodes de lyse bien connues, telles que les chocs thermiques ou osmotiques ou les lyses chimiques par des agents chaotropiques tels que les sels de guanidium (US 5,234,809).

- une étape de purification, permettant la séparation entre les acides nucléiques et les autres constituants cellulaires relargués dans l'étape de lyse. Cette étape permet généralement de concentrer les acides nucléiques, et peut être adapté à la purification d'ADN ou d'ARN. A titre d'exemple, on peut utiliser des particules magnétiques éventuellement revêtues d'oligonucléotides, par adsorption ou covalence (voir à ce sujet les brevets US 4,672,040 et US 5,750,338), et ainsi purifier les acides nucléiques qui se sont fixés sur ces particules magnétiques, par une étape de lavage. Cette étape de purification des acides nucléiques est particulièrement intéressante si l'on souhaite amplifier ultérieurement lesdits acides nucléiques. Un mode de réalisation particulièrement intéressant de ces particules magnétiques est décrit dans les demandes de brevet: WO-A-97/45202 et WO-A-99/35500. Un autre exemple intéressant de méthode de purification des acides nucléiques est l'utilisation de silice soit sous forme de colonne, soit sous forme de particules inertes (Boom R. et al., J. Clin. Microbiol., 1990, n°28(3), p. 495-503) ou magnétiques (Merck: MagPrep® Silica, Promega: MagneSil™ Paramagnetic particles). D'autres méthodes très répandues reposent sur des résines échangeuses d'ions en colonne ou en format particulière paramagnétique (Whatman: DEAE-Magarose) (Levison PR et al., J. Chromatography, 1998, p. 337-344). Une autre méthode très pertinente mais non exclusive pour l'invention est celle de l'adsorption sur support d'oxyde métallique (société Xtrana: matrice Xtra-Bind™).

Lorsque l'on souhaite extraire spécifiquement l'ADN d'un échantillon biologique, on peut notamment réaliser une extraction par du phénol, du chloroforme et de l'alcool pour éliminer les protéines et précipiter l'ADN avec de l'éthanol 100%. L'ADN peut alors être culoté par centrifugation, lavé et remis en solution.

Lorsque l'on souhaite extraire spécifiquement les ARN d'un échantillon biologique, on peut notamment réaliser une extraction par du phénol, du chloroforme et de

l'alcool pour éliminer les protéines et précipiter les ARN avec de l'éthanol 100%.

Les ARN peuvent alors être culoté par centrifugation, lavé et remis en solution.

Au sens de la présente invention, on entend par réactif spécifique, un réactif qui, lorsqu'il est mis en contact du matériel biologique tel que défini précédemment, se lie avec le matériel spécifique dudit gène cible. A titre indicatif, lorsque le réactif spécifique et le matériel biologique sont d'origine nucléique, la mise en contact du réactif spécifique et du matériel biologique permet l'hybridation du réactif spécifique avec le matériel spécifique du gène cible. Par hybridation, on entend le processus au cours duquel, dans des conditions appropriées, deux fragments nucléotidiques se lient avec des liaisons hydrogènes stables et spécifiques pour former un complexe double brin. Ces liaisons hydrogènes se forment entre les bases complémentaires Adénine (A) et thymine (T) (ou uracile (U)) (on parle de liaison A-T) ou entre les bases complémentaires Guanine (G) et cytosine (C) (on parle de liaison G-C). L'hybridation de deux fragments nucléotidiques peut être totale (on parle alors de fragments nucléotidiques ou de séquences complémentaires), c'est à dire que le complexe double brin obtenu lors de cette hybridation comprend uniquement des liaisons A-T et des liaisons C-G. Cette hybridation peut être partielle (on parle alors de fragments nucléotidiques ou de séquences suffisamment complémentaires), c'est à dire que le complexe double brin obtenu comprend des liaisons A-T et des liaisons C-G permettant de former le complexe double brin, mais également des bases non liées à une base complémentaire. L'hybridation entre deux fragments nucléotidiques dépend des conditions opératoires qui sont utilisées, et notamment de la stringence. La stringence est définie notamment en fonction de la composition en bases des deux fragments nucléotidiques, ainsi que par le degré de mésappariement entre deux fragments nucléotidiques. La stringence peut également être fonction des paramètres de la réaction, tels que la concentration et le type d'espèces ioniques présentes dans la solution d'hybridation, la nature et la concentration d'agents dénaturants et/ou la température d'hybridation. Toutes ces données sont bien connues et les conditions appropriées peuvent être déterminées par l'homme du métier. En général, selon la longueur des fragments nucléotidiques que l'on souhaite

hybrider, la température d'hybridation est comprise entre environ 20 et 70°C, en particulier entre 35 et 65°C dans une solution saline à une concentration d'environ 0,5 à 1 M. Une séquence, ou fragment nucléotidique, ou oligonucléotide, ou polynucléotide, est un enchaînement de motifs nucléotidiques assemblés entre eux par des liaisons ester phosphorique, caractérisé par la séquence informationnelle des acides nucléiques naturels, susceptibles de s'hybrider à un fragment nucléotidique, l'enchaînement pouvant contenir des monomères de structures différentes et être obtenu à partir d'une molécule d'acide nucléique naturelle et/ou par recombinaison génétique et/ou par synthèse chimique. Un motif est dérivé d'un monomère qui peut être un nucléotide naturel d'acide nucléique dont les éléments constitutifs sont un sucre, un groupement phosphate et une base azotée ; dans l'ADN le sucre est le désoxy-2-ribose, dans l'ARN le sucre est le ribose ; selon qu'il s'agisse de l'ADN ou l'ARN, la base azotée est choisie parmi l'adénine, la guanine, l'uracile, la cytosine, la thymine ; ou bien le monomère est un nucléotide modifié dans l'un au moins des trois éléments constitutifs ; à titre d'exemple, la modification peut intervenir soit au niveau des bases, avec des bases modifiées telles que l'inosine, la méthyl-5désoxycytidine, la désoxyuridine, la diméthylamino-5désoxyuridine, la diamino-2,6-purine, la bromo-5désoxyuridine ou toute autre base modifiée capable d'hybridation, soit au niveau du sucre, par exemple le remplacement d'au moins un désoxyribose par un polyamide (P.E. Nielsen et al, Science, 254, 1497-1500 (1991), soit encore au niveau du groupement phosphate, par exemple son remplacement par des esters notamment choisis parmi les diphosphates, alkyl- et aryl-phosphonates et phosphorothioates.

Selon un mode particulier de réalisation de l'invention, le réactif spécifique comprend au moins une amorce d'amplification. Au sens de la présente invention, on entend par amorce d'amplification, un fragment nucléotidique comprenant de 5 à 100 motifs nucléiques, préférentiellement de 15 à 30 motifs nucléiques permettant l'initiation d'une polymérisation enzymatique, telle que notamment une réaction d'amplification enzymatique. Selon un mode particulier de réalisation de l'invention, l'amorce d'amplification comprend une séquence

choisie parmi les SEQ ID N°38 à 41 et SEQ ID N°44 à 45. Par réaction d'amplification enzymatique, on entend un processus générant de multiples copies d'un fragment nucléotidique par l'action d'au moins une enzyme. De telles réactions d'amplification sont bien connues de l'homme du métier et on peut citer notamment les techniques suivantes :

- 5 - PCR (Polymerase Chain Reaction), telle que décrite dans les brevets US 4,683,195, US 4,683,202 et US 4,800,159,
- LCR (Ligase Chain Reaction), exposée par exemple dans la demande de brevet EP 0 201 184,
- RCR (Repair Chain Reaction), décrite dans la demande de brevet WO 90/01069,
- 10 - 3SR (Self Sustained Sequence Replication) avec la demande de brevet WO 90/06995,
- NASBA (Nucleic Acid Sequence-Based Amplification) avec la demande de brevet WO 91/02818, et
- TMA (Transcription Mediated Amplification) avec le brevet US 5,399,491.

15 Lorsque l'amplification enzymatique est une PCR, le réactif spécifique comprend au moins 2 amorces d'amplification, spécifiques d'un gène cible, et permettent l'amplification du matériel spécifique du gène cible. Le matériel spécifique du gène cible comprend alors préférentiellement un ADN complémentaire obtenu par transcription inverse d'ARN messager issu du gène cible (on parle alors d'ADNc spécifique du gène cible) ou un ARN

20 complémentaire obtenu par transcription des ADNc spécifique d'un gène cible (on parle alors d'ARNc spécifique du gène cible). Lorsque l'amplification enzymatique est une PCR réalisée après une réaction de transcription reverse, on parle de RT-PCR.

Selon un autre mode préféré de réalisation de l'invention, le réactif spécifique de l'étape b) comprend préférentiellement une sonde d'hybridation.

25 Par sonde d'hybridation, on entend un fragment nucléotidique comprenant de 5 à 100 motifs nucléiques, notamment de 10 à 35 motifs nucléiques, possédant une spécificité d'hybridation dans des conditions déterminées pour former un complexe d'hybridation avec le matériel spécifique d'un gène cible. Dans la présente invention, le matériel spécifique du gène cible peut être une séquence nucléotidique comprise dans un ARN messager issu du

gène cible (on parle alors d'ARNm spécifique du gène cible), une séquence nucléotidique comprise dans un ADN complémentaire obtenu par transcription inverse dudit ARN messager (on parle alors d'ADNc spécifique du gène cible), ou encore une séquence nucléotidique comprise dans un ARN complémentaire obtenu par transcription dudit

5 ADNc tel que décrit précédemment (on parlera alors d'ARNc spécifique du gène cible). La sonde d'hybridation peut comprendre un marqueur permettant sa détection. Par détection on entend soit une détection directe par une méthode physique, soit une détection indirecte par une méthode de détection à l'aide d'un marqueur. De nombreuses méthodes de détection existent pour la détection des acides nucléiques. [Voir par exemple Kricka et

10 al., Clinical Chemistry, 1999, n° 45(4), p.453-458 ou Keller G.H. et al., DNA Probes, 2nd Ed., Stockton Press, 1993, sections 5 et 6, p.173-249]. Par marqueur, on entend un traceur capable d'engendrer un signal que l'on peut détecter. Une liste non limitative de ces traceurs comprend les enzymes qui produisent un signal détectable par exemple par colorimétrie, fluorescence ou luminescence, comme la peroxydase de raifort, la

15 phosphatase alcaline, la beta galactosidase, la glucose-6-phosphate déshydrogénase; les chromophores comme les composés fluorescents, luminescents ou colorants ; les groupements à densité électronique détectables par microscopie électronique ou par leurs propriétés électriques comme la conductivité, par les méthodes d'ampérométrie ou de voltamétrie, ou par des mesures d'impédance ; les groupements détectables par des

20 méthodes optiques comme la diffraction, la résonance plasmon de surface, la variation d'angle de contact ou par des méthodes physiques comme la spectroscopie de force atomique, l'effet tunnel, etc. ; les molécules radioactives comme  $^{32}\text{P}$ ,  $^{35}\text{S}$  ou  $^{125}\text{I}$ .

Au sens de la présente invention, la sonde d'hybridation peut être une sonde dite de détection. Dans ce cas, la sonde dite de détection est marquée au moyen d'un marqueur tel

25 que défini précédemment. La sonde d'hybridation peut être également une sonde dite de capture. Dans ce cas, la sonde dite de capture est immobilisée ou immobilisable sur un support solide par tout moyen approprié, c'est-à-dire directement ou indirectement, par exemple par covalence ou adsorption. Comme support solide, on peut utiliser des matériaux de synthèse ou des matériaux naturels, éventuellement modifiés chimiquement,

notamment les polysaccharides tels que les matériaux à base de cellulose, par exemple du papier, des dérivés de cellulose tels que l'acétate de cellulose et la nitrocellulose ou le dextrane, des polymères, des copolymères, notamment à base de monomères du type styrène, des fibres naturelles telles que le coton, et des fibres synthétiques telles que le nylon ; des matériaux minéraux tels que la silice, le quartz, des verres, des céramiques ; des latex ; des particules magnétiques ; des dérivés métalliques, des gels etc. Le support solide peut être sous la forme d'une plaque de microtitration, d'une membrane comme décrit dans la demande WO-A-94/12670, d'une particule. On peut également immobiliser sur le support plusieurs sondes de capture différentes, chacune étant spécifique d'un gène cible.

En particulier, on peut utiliser comme support une biopuce sur laquelle peuvent être immobilisées un grand nombre de sondes. Par biopuce, on entend un support solide de dimension réduite où sont fixées une multitude de sondes de capture à des positions prédéterminées. Le concept de biopuce, ou puce à ADN, date du début des années 90. Il repose sur une technologie pluridisciplinaire intégrant la micro-électronique, la chimie des acides nucléiques, l'analyse d'images et l'informatique. Le principe de fonctionnement repose sur un fondement de la biologie moléculaire : le phénomène d'hybridation, c'est-à-dire l'appariement par complémentarité des bases de deux séquences d'ADN et/ou d'ARN. La méthode des biopuces repose sur l'emploi de sondes de capture fixées sur un support solide sur lesquelles on fait agir un échantillon de fragments nucléotidiques cibles marqués directement ou indirectement avec des fluorochromes. Les sondes de capture sont positionnées de manière spécifique sur le support ou puce et chaque hybridation donne une information particulière, en relation avec le fragment nucléotidique cible. Les informations obtenues sont cumulatives, et permettent par exemple de quantifier le niveau d'expression d'un gène ou de plusieurs gènes cibles. Pour analyser l'expression d'un gène cible, on peut alors réaliser une biopuce portant de très nombreuses sondes qui correspondent à tout ou partie du gène cible, qui est transcrit en ARNm. On hybride alors par exemple les ADNc ou les ARNc spécifiques d'un gène cible que l'on souhaite analyser sur des sondes de capture spécifique. Après hybridation, le support ou puce est lavé(e), et les complexes ADNc ou ARNc marquées / sondes de capture sont révélés par un ligand de forte affinité



lié par exemple à un marqueur de type fluorochrome. La fluorescence est lue par exemple par un scanner et l'analyse de la fluorescence est traitée par informatique. On peut citer à titre indicatif, les puces à ADN mises au point par la société Affymetrix ("Accessing Genetic Information with High-Density DNA arrays", M. Chee et al., Science, 1996, 274, 610-614. "Light-generated oligonucleotide arrays for rapide DNA sequence analysis", A. Caviani Pease et al., Proc. Natl. Acad. Sci. USA, 1994, 91, 5022-5026), pour les diagnostics moléculaires. Dans cette technologie, les sondes de capture sont généralement de tailles réduites, autour de 25 nucléotides. D'autres exemples de biopuces sont donnés dans les publications de G. Ramsay, Nature Biotechnology, 1998, n°16, p. 40-44 ; F. Ginot, Human Mutation, 1997, n°10, p.1-10; J. Cheng et al, Molecular diagnosis, 1996, n°1(3), p.183-200 ; T. Livache et al, Nucleic Acids Research, 1994, n° 22(15), p. 2915-2921 ; J. Cheng et al, Nature Biotechnology, 1998, n° 16, p. 541-546 ou dans les brevets US-A-4,981,783, US-A-5,700,637, US-A-5,445,934, US-A-5,744,305 et US-A-5,807,522. La caractéristique principale du support solide doit être de conserver les caractéristiques d'hybridation des sondes de capture sur les fragments nucléotidiques cibles tout en générant un bruit de fond minimum pour la méthode de détection.

Pour l'immobilisation des sondes sur le support, on distingue trois grands types de fabrication.

Il y a, tout d'abord, une première technique qui consiste en un dépôt de sondes pré-synthétisées. La fixation des sondes se fait par transfert direct, au moyen de micropipettes, de micro-pointes ou par un dispositif de type jet d'encre. Cette technique permet la fixation de sondes de taille allant de quelques bases (5 à 10) jusqu'à des tailles relativement importantes de 60 bases (impression) à quelques centaines de bases (micro-déposition) :

L'impression est une adaptation du procédé utilisé par les imprimantes à jet d'encre. Elle repose sur la propulsion de très petites sphères de fluide (volume <1 nl) et à un rythme pouvant atteindre 4000 gouttes/secondes. L'impression n'implique aucun contact entre le système libérant le fluide et la surface sur laquelle il est déposé.

La micro-déposition consiste à fixer des sondes longues de quelques dizaines à plusieurs centaines de bases à la surface d'une lame de verre. Ces sondes sont généralement extraites de bases de données et se présentent sous forme de produits amplifiés et purifiés. Cette technique permet de réaliser des puces dénommées microarrays  
5 portant environ dix mille spots, dit zones de reconnaissance, d'ADN sur une surface d'un peu moins de 4 cm<sup>2</sup>. Il ne faut toutefois pas oublier l'emploi de membranes de Nylon, dites « macroarrays », qui portent des produits amplifiés, généralement par PCR, avec un diamètre de 0,5 à 1 mm et dont la densité maximale est de 25 spots/cm<sup>2</sup>. Cette technique très flexible est utilisée par de nombreux laboratoires. Dans la présente invention, cette  
10 dernière technique est considérée comme faisant partie des biopuces. On peut toutefois déposer en fond de plaque de microtitration un certain volume d'échantillon dans chaque puits, comme c'est le cas dans les demandes de brevet WO-A-00/71750 et FR 00/14896, ou déposer au fond d'une même boîte de Pétri un certain nombre de gouttes séparées les unes des autres, selon une autre demande de brevet FR00/14691.

15 La deuxième technique de fixation des sondes sur le support ou puce est appelée la synthèse in situ. Cette technique aboutit à l'élaboration de sondes courtes directement à la surface de la puce. Elle repose sur la synthèse d'oligonucléotides in situ (voir notamment les demandes de brevet WO 89/10977 et WO 90/03382), et est fondée sur le procédé des synthétiseurs d'oligonucléotides. Elle consiste à déplacer une chambre de réactions, où se  
20 déroule la réaction d'élongation d'oligonucléotides, le long de la surface de verre.

Enfin, la troisième technique est appelée la photolithographie, qui est un procédé à l'origine des biopuces développées par Affymetrix. Il s'agit également d'une synthèse in situ. La photolithographie est dérivée des techniques des microprocesseurs. La surface de la puce est modifiée par la fixation de groupements chimiques photolabiles pouvant être  
25 activés par la lumière. Une fois illuminés, ces groupes sont susceptibles de réagir avec l'extrémité 3' d'un oligonucléotide. En protégeant cette surface par des masques de formes définies, on peut illuminer et donc activer sélectivement des zones de la puce où l'on souhaite fixer l'un ou l'autre des quatre nucléotides. L'utilisation successive de masques différents permet d'alterner des cycles de protection/réaction et donc de réaliser les sondes

d'oligonucléotides sur des spots d'environ quelques dizaines de micromètre carré ( $\mu\text{m}^2$ ). Cette résolution permet de créer jusqu'à plusieurs centaines de milliers de spots sur une surface de quelques centimètres carré ( $\text{cm}^2$ ). La photolithographie présente des avantages : massivement parallèle, elle permet de créer une puce de N-mères en seulement  $4 \times N$  cycles. Toutes ces techniques sont bien entendues utilisables avec la présente invention. Selon un mode préféré de réalisation de l'invention, le au moins un réactif spécifique de l'étape b) définie précédemment comprend au moins une sonde d'hybridation, qui est préférentiellement immobilisée sur un support. Ce support est préférentiellement une biopuce telle que définie précédemment.

10

Lors de l'étape c) la détermination de l'expression d'un gène cible peut être réalisée par tous les protocoles connus de l'homme du métier.

D'une manière générale, l'expression d'un gène cible peut être analysée par la détection des ARNm (ARN messagers) qui sont transcrits du gène cible à un instant donné ou par la détection des protéines issues de ces ARNm.

15

L'invention concerne préférentiellement la détermination de l'expression d'un gène cible par la détection des ARNm issus de ce gène cible selon tous les protocoles bien connus de l'homme du métier. Selon un mode particulier de réalisation de l'invention, on détermine simultanément l'expression de plusieurs gènes cibles, par la détection de plusieurs ARNm différents, chaque ARNm étant issu d'un gène cible.

20

Lorsque le réactif spécifique comprend au moins une amorce d'amplification, on peut, lors de l'étape c) du procédé selon l'invention, déterminer l'expression d'un gène cible de la manière suivante:

1) après avoir extrait comme matériel biologique, les ARN totaux (comprenant les ARN de transfert (ARNt), les ARN ribosomaux (ARNr) et les ARN messagers (ARNm)) d'un échantillon biologique tel que présenté précédemment, on réalise une étape de transcription reverse afin d'obtenir les ADN complémentaires (ou ADNc) desdits ARNm. A titre indicatif, cette réaction de transcription reverse peut être réalisée à l'aide d'une enzyme reverse transcriptase qui permet d'obtenir, à partir d'un fragment d'ARN, un

25

fragment d'ADN complémentaire. On peut utiliser notamment l'enzyme reverse transcriptase provenant de l'AMV (Avian Myoblastosis Virus) ou de MMLV (Moloney Murine Leukaemia Virus). Lorsque l'on souhaite plus particulièrement obtenir uniquement les ADNc des ARNm, on réalise cette étape de transcription reverse en présence de  
5 fragments nucléotidiques comprenant uniquement des bases thymine (polyT), qui s'hybrident par complémentarité sur la séquence polyA des ARNm afin de former un complexe polyT-polyA qui sert alors de point de départ à la réaction de transcription reverse réalisée par l'enzyme reverse transcriptase. On obtient alors des ADNc complémentaires des ARNm issus d'un gène cible (ADNc spécifique du gène cible) et des  
10 ADNc complémentaires des ARNm issus d'autres gènes que le gène cible (ADNc non spécifique du gène cible).

2) on met en contact la ou les amorces d'amplification spécifiques d'un gène cible avec les ADNc spécifique du gène cible et les ADNc non spécifique du gène cible. La ou les amorces d'amplification spécifiques d'un gène cible s'hybrident avec les ADNc  
15 spécifique du gène cible et on amplifie spécifiquement une région prédéterminée, de longueur connue, des ADNc provenant des ARNm issus du gène cible. Les ADNc non spécifiques du gène cible ne sont pas amplifiés, alors qu'on obtient alors une grande  
quantité d'ADNc spécifiques du gène cible. Au sens de la présente invention, on parle indifféremment d' « ADNc spécifiques du gène cible » ou d' « ADNc provenant des  
20 ARNm issus du gène cible ». Cette étape peut être réalisée notamment par une réaction d'amplification de type PCR ou par toute autre technique d'amplification telle que définie précédemment. En PCR, on peut également amplifier simultanément plusieurs ADNc différents, chacun étant spécifique de différent gène cible par l'utilisation de plusieurs couples d'amorces d'amplification différentes, chacune étant spécifique d'un gène cible: on  
25 parle alors d'amplification en multiplex.

3) on détermine l'expression du gène cible en détectant et quantifiant les ADNc spécifiques du gène cible obtenus lors de l'étape 2) ci dessus. Cette détection peut être réalisée après migration par électrophorèse des ADNc spécifiques du gène cible en fonction de leur taille. Le gel et le milieu de migration peuvent comprendre du bromure

d'éthydium afin de permettre la détection directe des ADNc spécifiques du gène cible lorsque le gel est placé, après un temps de migration donné, sur une table lumineuse à rayons UV (ultra violet) par l'émission d'un signal lumineux. Ce signal est d'autant plus lumineux que la quantité des ADNc spécifique du gène cible est importante. Ces techniques d'électrophorèse sont bien connues de l'homme du métier. Les ADNc spécifiques du gène cible peuvent également être détectés et quantifiés par l'utilisation d'une gamme de quantification obtenue par une réaction d'amplification conduite jusqu'à saturation. Afin de tenir compte de la variabilité d'efficacité enzymatique qui peut être observée lors des différentes étapes (transcription reverse, PCR...), on peut normaliser l'expression d'un gène cible de différents groupes de patients, par la détermination simultanée de l'expression d'un gène dit de ménage, dont l'expression est similaire chez les différents groupes de patients. En réalisant un rapport entre l'expression du gène cible et l'expression du gène de ménage, c'est à dire en réalisant un rapport entre la quantité d'ADNc spécifiques du gène cible, et la quantité d'ADNc spécifiques du gène de ménage, on corrige ainsi toute variabilité entre les différentes expérimentations. L'homme du métier pourra se référer notamment aux publications suivantes : Bustin SA *Journal of molecular endocrinology*, 2002, 29 : 23-39 ; Giulietti A *Methods*, 2001, 25 : 386-401.

Lorsque le réactif spécifique comprend au moins une sonde d'hybridation, on peut déterminer l'expression d'un gène cible de la manière suivante:

1) après avoir extrait, comme matériel biologique, les ARN totaux d'un échantillon biologique tel que présenté précédemment, on réalise une étape de transcription reverse, telle que décrite précédemment afin des ADNc complémentaires des ARNm issus d'un gène cible (ADNc spécifique du gène cible) et des ADNc complémentaires des ARNm issus d'autres gènes que le gène cible (ADNc non spécifique du gène cible).

2) on met en contact tous les ADNc avec un support, sur lequel sont immobilisées des sondes de capture spécifiques du gène cible dont on souhaite analyser l'expression, afin de réaliser une réaction d'hybridation entre les ADNc spécifiques du gène cible et les sondes de capture, les ADNc non spécifiques du gène cible ne s'hybrident pas sur les

sondes de capture. La réaction d'hybridation peut être réalisée sur un support solide qui inclut tous les matériaux tels qu'indiqué précédemment. Selon un mode préféré de réalisation, la sonde d'hybridation est immobilisée sur un support. Préférentiellement, le support est une biopuce. La réaction d'hybridation peut être précédée d'une étape d'amplification enzymatique des ADNc spécifique du gène cible telle que décrite précédemment pour obtenir une grande quantité d'ADNc spécifiques du gène cible et augmenter la probabilité qu'un ADNc spécifique d'un gène cible s'hybride sur une sonde de capture spécifique du gène cible. La réaction d'hybridation peut également être précédée d'une étape de marquage et/ou de clivage des ADNc spécifiques du gène cible telle que décrite précédemment, par exemple en utilisant un désoxyribonucléotide triphosphate marqué pour la réaction d'amplification. Le clivage peut être réalisé notamment par l'action de l'imidazole et de chlorure de manganèse. L'ADNc spécifique du gène cible peut aussi être marqué après l'étape d'amplification, par exemple en hybridant une sonde marquée selon la technique d'hybridation sandwich décrite dans le document WO-A-91/19812. D'autres modes particuliers préférentiels de marquage et/ou clivage d'acides nucléiques sont décrits dans les demandes WO 99/65926, WO 01/44507, WO 01/44506, WO 02/090584, WO 02/090319.

3) on réalise ensuite une étape de détection de la réaction d'hybridation. La détection peut être réalisée par la mise en contact du support sur lequel sont hybridés les sondes de capture spécifique du gène cible avec les ADNc spécifiques du gène cible avec une sonde dite de détection, marquée par un marqueur, et on détecte le signal émis par le marqueur. Lorsque l'ADNc spécifique du gène cible a été préalablement marqué par un marqueur, on détecte directement le signal émis par le marqueur.

Lorsque le au moins un réactif spécifique mis en contact l'étape b) du procédé selon l'invention comprend au moins une sonde d'hybridation, on peut également déterminer l'expression d'un gène cible de la manière suivante:

1) après avoir extrait, comme matériel biologique, les ARN totaux d'un échantillon biologique telle que présentée précédemment, on réalise une étape de transcription reverse,

telle que décrite précédemment afin d'obtenir les ADNc des ARNm du matériel biologique. On réalise ensuite la polymérisation de l'ARN complémentaire du ADNc par l'utilisation d'une enzyme polymérase de type T7 polymérase qui fonctionnent sous la dépendance d'un promoteur et qui permettent d'obtenir, à partir d'une matrice d'ADN, l'ARN  
5 complémentaire. On obtient alors les ARNc des ADNc des ARNm spécifiques du gène cible (on parle alors d'ARNc spécifique du gène cible) et les ARNc des ADNc des ARNm non spécifiques du gène cible.

2) on met en contact tous les ARNc avec un support, sur lequel sont immobilisées des sondes de capture spécifiques du gène cible dont on souhaite analyser l'expression,  
10 afin de réaliser une réaction d'hybridation entre les ARNc spécifiques du gène cible et les sondes de capture, les ARNc non spécifiques du gène cible ne s'hybridant pas sur les sondes de capture. Lorsque l'on souhaite analyser simultanément l'expression de plusieurs gènes cibles, on peut immobiliser sur le support plusieurs sondes de capture différentes, chacune étant spécifique d'un gène cible. La réaction d'hybridation peut également être  
15 précédée d'une étape de marquage et/ou de clivage des ARNc spécifiques du gène cible telles que décrites précédemment.

3) on réalise ensuite une étape de détection de la réaction d'hybridation. La détection peut être réalisée par la mise en contact du support sur lequel sont hybridées les sondes de capture spécifiques du gène cible avec l'ARNc spécifique du gène cible avec  
20 une sonde dite de détection, marquée par un marqueur, et on détecte le signal émis par le marqueur. Lorsque l'ARNc spécifiques du gène cible a été préalablement marqué par un marqueur, on détecte directement le signal émis par le marqueur. L'utilisation d'ARNc est particulièrement avantageux lorsqu'on utilise un support de type biopuce sur lequel est hybridés un grand nombre de sondes.

25

L'analyse de l'expression d'un gène cible choisi parmi l'une quelconque des SEQ ID N°1 à 37 permet alors de disposer d'un outil pour le pronostic du neuroblastome. On peut par exemple analyser l'expression d'un gène cible chez un patient dont on ne connaît pas le pronostic, et comparer avec des valeurs d'expression moyenne connues du gène cible de

patients de bon pronostic et des valeurs d'expression moyenne connues du gène cible de patients de mauvais pronostic. Ceci permet de déterminer si le patient est de bon ou de mauvais pronostic afin de lui proposer un traitement adapté.

- 5 Selon un mode préféré de réalisation de l'invention, lors de l'étape b) on met en contact le matériel biologique avec au moins 37 réactifs spécifiques choisis parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant l'une quelconque des SEQ ID N° 1 à 37 et on détermine, lors de l'étape c, l'expression d'au moins 37 desdits gènes cibles.
- 10 Selon un autre mode préféré de réalisation, lors de l'étape b) on met en contact le matériel biologique avec au moins 2, au moins 3, au moins 4, au moins 5, au moins 6, au moins 7, au moins 8, au moins 9, au moins 10, au moins 11, au moins 12, au moins 13, au moins 14, au moins 15, au moins 16, au moins 17, au moins 18, au moins 19, au moins 20, au moins 21, au moins 22, au moins 23, au moins 24, au moins 25, au moins 26, au moins 27, au moins 28, au moins 29, au moins 30, au moins 31, au moins 32, au moins 33, au moins 34, au moins 35, ou au moins 36 réactifs spécifiques choisis parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant l'une quelconque des SEQ ID N° 1 à 37 et on détermine, lors de l'étape c, l'expression d'au moins au moins 2, au moins 3, au moins 4, au moins 5, au moins 6, au moins 7, au moins 8, au moins 9, au moins 10, au moins 11, au moins 12, au moins 13, au moins 14, au moins 15, au moins 16, au moins 17, au moins 18, au moins 19, au moins 20, au moins 21, au moins 22, au moins 23, au moins 24, au moins 25, au moins 26, au moins 27, au moins 28, au moins 29, au moins 30, au moins 31, au moins 32, au moins 33, au moins 34, au moins 35, ou au moins 36 desdits gènes cibles.
- 25 Selon un autre mode préféré de réalisation, lors de l'étape b) on met en contact le matériel biologique avec au moins 19 réactifs spécifiques choisis parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant la SEQ ID N°1 ; SEQ ID N°2; SEQ ID N°3; SEQ ID N°7; SEQ ID N°8; SEQ ID N°9; SEQ ID N°10; SEQ ID N°14; SEQ ID N°16; SEQ ID N°20; SEQ ID N°21; SEQ ID N°22; SEQ ID N°25; SEQ ID N°27;



SEQ ID N°29; SEQ ID N°31; SEQ ID N°34; SEQ ID N°36 ou SEQ ID N°37 on détermine, lors de l'étape c) l'expression d'au moins 19 desdits gènes cibles.

Selon un autre mode préféré de réalisation, lors de l'étape b) on met en contact le matériel biologique avec au moins 2, au moins 3, au moins 4, au moins 5, au moins 6, au moins 7,  
5 au moins 8, au moins 9, au moins 10, au moins 11, au moins 12, au moins 13, au moins 14, au moins 15, au moins 16, au moins 17, au moins 18, ou au moins 19 réactifs spécifiques choisis parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant la SEQ ID N°1 ; SEQ ID N°2; SEQ ID N°3; SEQ ID N°7; SEQ ID N°8; SEQ ID N°9; SEQ ID N°10; SEQ ID N°14; SEQ ID N°16; SEQ ID N°20; SEQ  
10 ID N°21; SEQ ID N°22; SEQ ID N°25; SEQ ID N°27; SEQ ID N°29; SEQ ID N°31; SEQ ID N°34; SEQ ID N°36 ou SEQ ID N°37 on détermine, lors de l'étape c) l'expression d'au moins 2, au moins 3, au moins 4, au moins 5, au moins 6, au moins 7, au moins 8, au moins 9, au moins 10, au moins 11, au moins 12, au moins 13, au moins 14, au moins 15, au moins 16, au moins 17, au moins 18, ou au moins 19 desdits gènes cibles.

Selon un autre mode préféré de réalisation, lors de l'étape b) on met en contact le matériel biologique avec au moins 16 réactifs spécifiques choisis parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant la SEQ ID N°1 ;  
15 SEQ ID N°2; SEQ ID N°3; SEQ ID N°7; SEQ ID N°8; SEQ ID N°9; SEQ ID N°10; SEQ ID N°20; SEQ ID N°21; SEQ ID N°22; SEQ ID N°25; SEQ ID N°29; SEQ ID N°31; SEQ ID N°34; SEQ ID N°36 ou SEQ ID N°37 on détermine, lors de l'étape c) l'expression d'au moins 16 desdits gènes cibles.

Selon un autre mode préféré de réalisation, lors de l'étape b) on met en contact le matériel biologique avec au moins 12 réactifs spécifiques choisis parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant la SEQ ID N°2;  
25 SEQ ID N°3; SEQ ID N°7; SEQ ID N°8; SEQ ID N°10; SEQ ID N°20; SEQ ID N°22; SEQ ID N°25; SEQ ID N°29; SEQ ID N°31; SEQ ID N°34; ou SEQ ID N°37 et on détermine, lors de l'étape c) l'expression d'au moins 12 desdits gènes cibles.

Selon un autre mode préféré de réalisation, lors de l'étape b) on met en contact le matériel biologique avec au moins 9 réactifs spécifiques choisis parmi les réactifs

spécifiques des gènes cibles présentant une séquence nucléique ayant la SEQ ID N°2; SEQ ID N°3; SEQ ID N°7; SEQ ID N°8; SEQ ID N°10; SEQ ID N°22; SEQ ID N°25; SEQ ID N°29; SEQ ID N°34; et on détermine, lors de l'étape c) l'expression d'au moins 9 desdits gènes cibles.

5 L'utilisation d'un panel de gènes restreint est particulièrement adapté pour obtenir un outil de pronostic. En effet, l'analyse de l'expression d'une dizaine de gènes ne nécessite pas la fabrication à façon de puces à ADN, et peut être mise en œuvre directement par des techniques de PCR ou de NASBA, ce qui présente un atout économique important et une mise en œuvre simplifiée.

10 Les figures ci-jointes sont données à titre d'exemples explicatifs et n'ont aucun caractère limitatif. Elles permettront de mieux comprendre l'invention.

La figure 1 présente un dendrogramme obtenu à partir de 23 échantillons de tumeurs issus de patients de bon pronostic (BP) ou de mauvais pronostic (MP), et l'utilisation d'un panel de 40 sondes permettant l'analyse de l'expression des 37 gènes  
15 présentés précédemment dans le tableau 1. On retrouve sur ce dendrogramme 23 colonnes correspondant aux 23 échantillons de tumeurs, et 40 lignes correspondant aux 40 sondes utilisées pour l'analyse de l'expression des 37 gènes. Les échantillons de tumeurs, ainsi que les gènes ayant un profil d'expression comparable, mis en évidence par une corrélation de type Pearson, ont été placés côte à côte. Les échantillons de tumeurs ont été classés selon  
20 la méthode de moyenne non-pondérée (Spotfire Decision Site for Functional Genomics V7.1, manual) alors que les gènes ont été classés selon la valeur moyenne d'expression obtenue dans l'ensemble des échantillons. Le niveau d'expression de chaque gène, calculé par le logiciel Microarray Suite (MAS5.0, Affymetrix) est représenté par différents niveaux de couleur. Ainsi, la couleur blanche correspond à un faible niveau d'expression, la couleur  
25 grise correspond à un niveau d'expression intermédiaire, alors que la couleur noire correspond à un fort niveau d'expression. La longueur des branches du dendrogramme est corrélée au profil d'expression et la ligne en pointillée qui divise le dendrogramme permet de distinguer deux groupes de patients : un premier groupe de patients de mauvais pronostic "MP" et un deuxième groupe de patients de bon pronostic "BP". Les six tumeurs "MP-

test" et «BP-test» sont des tumeurs qui ont été analysées «en aveugle», c'est à dire sans connaître leur pronostic.

La figure 2 présente un dendogramme obtenu à partir de 23 échantillons de tumeurs issus de patients de bon pronostic ou de mauvais pronostic, et l'analyse de l'expression de 19 gènes. Ce dendogramme a été obtenu comparablement à ce qui est décrit pour la figure 1.

La figure 3 présente un dendogramme obtenu à partir de 23 échantillons de tumeurs issus de patients de bon pronostic ou de mauvais pronostic, et l'analyse de l'expression de 16 gènes. Ce dendogramme a été obtenu comparablement à ce qui est décrit pour la figure 1.

La figure 4 présente un dendogramme obtenu à partir de 23 échantillons de tumeurs issus de patients de bon pronostic ou de mauvais pronostic, et l'analyse de l'expression de 12 gènes. Ce dendogramme a été obtenu comparablement à ce qui est décrit pour la figure 1.

La figure 5 présente un dendogramme obtenu à partir de 23 échantillons de tumeurs issus de patients de bon pronostic ou de mauvais pronostic, et l'analyse de l'expression de 9 gènes. Ce dendogramme a été obtenu comparablement à ce qui est décrit pour la figure 1.

Les exemples suivants sont donnés à titre illustratif et n'ont aucun caractère limitatif. Ils permettront de mieux comprendre l'invention.

#### **Exemple 1 : Recherche d'un profil d'expression pour le pronostic du neuroblastome**

Caractéristiques des échantillons biologiques (tumeurs localisées ou ponctions de moelles osseuses) : 23 échantillons de neuroblastome, obtenus auprès du Centre Léon Bérard (CLB) de Lyon, France, ont été utilisées dans cette étude. Ces échantillons de neuroblastome ont été prélevés préalablement à tout traitement thérapeutique. Chaque tumeur a été classée suivant la classification INSS (International Neuroblastoma Staging

System; Brodeur et al; (1993) *J. Clin. Oncol.* 11, 1466-77). On distinguait alors 12 tumeurs de stade 1/2, 4 tumeurs de stade 4s et 7 échantillons de stade 4. (2 ponctions tumorales, 1 biopsie, 4 ponctions médullaires massivement envahies. L'analyse histochimique montrait dans les tumeurs localisées la présence d'environ 80% de cellules tumorales. L'analyse immunocytochimique montrait également dans les ponctions de moelle osseuse la présence d'environ 80% de cellules tumorales. L'âge médian des patients au moment du diagnostic du neuroblastome était de 10 mois et demi, et 5 patients sont décédés au cours de la période de suivi médian de 75 mois. Les patients ayant décédé au cours de l'étude, et les patients présentant un neuroblastome de stade IV étaient qualifiés de patients de mauvais pronostic (MP), alors que les patients en vie, ayant développé un neuroblastome de stade 1, 2 et 4s étaient qualifiés de patients de bon pronostic (BP) (qualification selon Brodeur, 2003, *Nat Rev Cancer*, 203-216). Cette analyse a ainsi été réalisée sur 8 patients MP et 15 patients BP.

Extraction du matériel biologique (ARN totaux) de l'échantillon biologique : les ARN totaux ont été extraits de chaque tumeur ou ponction de moelle osseuse selon un protocole bien connu de l'homme du métier (Voir notamment Ausubel et al (1997), *Current protocols in Molecular Biology*, Volume 1, John Wiley and Sons, New York). Pour cela, chaque échantillon biologique a été homogénéisé dans 1 ml de Trizol (Invitrogen, Cergy Pointoise, France), et traité avec 300 µl de chloroforme afin d'éliminer tout contaminant protéique et lipophile. Les ARN totaux ont ensuite été précipités avec 750 µl d'isopropanol, lavés deux fois avec une solution à 80 % en ethanol (vol/vol) et remis en solution dans de l'eau DEPC. Les ARN totaux ont ensuite été purifiés sur colonne Qiagen RNeasy (Qiagen, Hilden, Germany) conformément aux instructions du fabricant, à l'exception de l'élution finale qui a été réalisée dans 200 µl d'eau RNase-free après 1 min d'incubation à 65°C. Préalablement à l'étape de transcription reverse, une étape de précipitation par de l'acétate d'ammonium (0,5 vol, 7.5M) et de l'éthanol (2,5 vol) a été réalisée pour garantir la purification des ARN totaux. La qualité des ARN totaux a été analysée par le bio analyseur AGILENT 2100 (Agilent Technologies, Waldbronn,

Germany). Les ARN totaux comprennent les ARN de transfert, les ARN messagers (ARNm) et les ARN ribosomaux.

Synthèse d'ADNc, obtention des ARNc et marquage des ARNc et quantification : Afin

d'analyser l'expression des gènes cibles selon l'invention, les ADN complémentaires (ADNc) des ARNm contenus dans les ARN totaux tels que purifiés ci dessus, ont été  
5 obtenus à partir de 10 µg d'ARN totaux par l'utilisation de 400 unités de l'enzyme de transcription reverse SuperScriptII (Invitrogen) et 100 pmol d'amorce poly-T contenant le promoteur de la T7 promotor (T7-oligo(dT)24-primer, Proligo, Paris, France). Les ADNc ainsi obtenus ont ensuite été extraits avec du phénol/chloroforme, et précipités tels que  
10 décrit précédemment par de l'acétate d'ammonium et de l'éthanol, et remis en solution dans 24 µl d'eau DEPC. Un volume de 20 µl de cette solution purifiée d'ADNc a fait l'objet ensuite d'une transcription *in vitro* par l'utilisation d'une ARN polymérase T7 qui reconnaît spécifiquement le promoteur de la T7 polymérase tel que mentionné ci dessus. Cette transcription permet d'obtenir l'ARNc de l'ADNc. Cette transcription a été réalisée  
15 par l'utilisation d'un kit Bioarray High Yield RNA Transcript Labeling Kit (Enzo Diagnostics, Farmingdale, NY), qui permet non seulement d'obtenir l'ARNc mais également l'incorporation de bases cytidine et uridine biotinylées lors de la synthèse de l'ARNc.

Les ARNc purifiés ont ensuite été quantifiés par spectrophotométrie, et la solution d'ARNc  
20 a été ajustée à une concentration de 1 µg/µl d'ARNc. L'étape de clivage de ces ARNc a ensuite été réalisée à 94°C pendant 35 min, par l'utilisation d'un tampon de fragmentation (40 mM de Tris acétate, pH 8,1, 100 mM d'acétate de potassium, 30 mM d'acétate de magnésium) afin de provoquer l'hydrolyse des ARNc et obtenir des fragments de 35 à 200 bp. Le succès d'une telle fragmentation a été vérifié par une électrophorèse sur gel  
25 d'agarose 1,5%).

Mise en évidence d'un profil d'expression différentiel entre les patients BP et MP

L'expression d'environ 10 000 gènes a été analysée et comparée entre les patients BP et MP. Pour cela, 10 µg d'ARNc fragmentés issus de chaque échantillon ont été ajoutés à un tampon d'hybridation (Affymetrix) et 200 µl de cette solution ont été mis en contact

pendant 16 h à 45°C sur une puce d'expression (Human Genome U95Av2 GeneChip® (Affymetrix), qui comporte 12 625 groupes de sondes représentant environ 10 000 gènes selon le protocole d'Affymetrix tel que décrit sur le site internet d'Affymetrix (voir notamment à l'adresse suivante

5 [http://www.affymetrix.com/support/downloads/manuals/expression\\_s2\\_manual.pdf](http://www.affymetrix.com/support/downloads/manuals/expression_s2_manual.pdf)).

Afin d'enregistrer les meilleures performances d'hybridation et de lavage, des ARN qualifiés de « contrôle » biotinylés (bioB, bioC, bioD et cre) et des oligonucléotides (oligo B2) ont également été inclus dans le tampon d'hybridation. Après l'étape d'hybridation, la solution d'ARNc biotinylée et hybridée sur la puce, a été révélée par l'utilisation d'une  
10 solution de streptavidine-phycoerythrine et le signal a été amplifié par l'utilisation d'anticorps anti-streptavidine. L'hybridation a été réalisée dans une étuve d'hybridation « GeneChip Hybridisation oven » (Affymetrix), et le protocole Euk GE-WS2 du protocole d'Affymetrix a été suivi. Les étapes de lavage et de révélation ont été réalisées sur une station « Fluidics Station 400 » (Affymetrix). Chaque puce U95Av2 a ensuite été analysée  
15 sur un scanner Agilent G2500A GeneArray Scanner à une résolution de 3 microns afin de repérer les zones hybridées sur la puce. Ce scanner permet la détection du signal émis par les molécules fluorescentes après excitation par un laser argon en utilisant la technique du microscope à épifluorescence. On obtient ainsi pour chaque position, un signal proportionnel à la quantité de ARNc fixés. Le signal a ensuite été analysé par le logiciel  
20 Microarray Suite 5.0 software (MAS5.0, Affymetrix).

Afin de prévenir les variations obtenues par l'utilisation de différentes puces, il a été réalisé une approche de normalisation globale utilisant le logiciel MAS5.0 (Affymetrix), qui permet de convertir les données brutes obtenues pour chaque puce en un signal moyen d'une intensité de 500. Les résultats obtenus sur une puce peuvent alors être comparés aux  
25 résultats obtenus sur une autre puce. Le logiciel MAS5.0 permettait aussi d'inclure un algorithme statistique pour considérer si un gène était exprimé ou non. Chaque gène représenté sur la puce U95Av2 était couvert par 16 à 20 couples de sondes de 25 oligonucléotides. Par couple de sondes, on entend une première sonde qui s'hybridait parfaitement (on parle alors de sondes PM ou perfect match) avec un des ARNc issus

d'un gène cible, et une deuxième sonde, identique à la première sonde à l'exception d'un mésappariement (on parle alors de sonde MM ou mismatched) au centre de la sonde. Chaque sonde MM servait à estimer le bruit de fond correspondant à une hybridation entre deux fragments nucléotidiques de séquence non complémentaire. (Affymetrix technical note

5 "Statistical Algorithms Reference Guide"; Lipshutz, et al (1999) Nat. Genet. 1 Suppl., 20-24). Deux tumeurs de stade IV présentant un faible pourcentage de gènes exprimés, du à un biais soit dans la qualité des ARNc soit dans l'étape d'hybridation, ont été exclues de l'analyse. Les 23 échantillons restant montraient une moyenne de 48 % de gènes exprimés. L'analyse des données d'expression a été réalisée par le logiciel Microsoft Excel, le logiciel

10 Spotfire Decision Site for Functionnal Genomics V7.1 (Spotfire AB, Gothenburg, Sweden), ainsi que le module PAM (Prediction Analysis in Microarrays) du logiciel de statistiques R (Ihaka & Gentleman (1996) Journal of Computational and Graphical Statistics 5, 299-314. ; Tibshirani, et al (2002) Proc. Natl. Acad. Sci. 99, 6567-6572).

A partir des 12625 groupes de sondes, représentant environ 10 000 gènes, de la puce, les

15 inventeurs ont sélectionné les gènes pertinents qui étaient corrélés à un mauvais pronostic du neuroblastome.

Pour cela, une première étape a consisté à exclure les gènes présentant un niveau d'expression comparable entre tous les groupes de patients [Tibshirani, et al Proc. Natl. Acad. Sci. 99, 6567-6572]. Les gènes non exprimés chez l'ensemble des patients ont

20 également été exclus (logiciel MAS5.0). Enfin, certains gènes ont été exclus si la moyenne d'expression des 2 groupes (patients de bon pronostic et patient de mauvais pronostic) était inférieur à 500 ou si le rapport des moyennes d'expression entre les patients de mauvais et de bon pronostics étaient compris entre 0,7 et 1,3.

L'expression des 1488 gènes restants a ensuite été analysée (algorithme PAM, Tibshirani,

25 R., Hastie, T., Narasimhan, B. and Chu, G. (2002) Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc. Natl. Acad. Sci. 99, 6567-6572).

Résultats obtenus: Dans un premier temps, 37 gènes permettant de différencier les patients de bon et de mauvais pronostic ont été identifiés. L'augmentation ou la diminution

d'expression de chacun de ces gènes, observée chez les patients de mauvais pronostic par rapport aux patients de bons pronostics est indiquée dans le tableau 2.

**Tableau 2 - liste des 37 gènes exprimés différemment dans les neuroblastomes patients**

**BP et MP**

SEQ ID N°	Description de la séquence	N° Genbank	Expression MP vs BP
1	Flap structure-specific endonuclease 1	NM_004111	augmentée
2	Ubiquitin-conjugating enzyme E2C	NM_007019	augmentée
3	Insulin-like growth factor binding protein 7(MAC25)	NM_001553	diminuée
4	Collagen type I, alpha 2 chain	NM_000089	diminuée
5	Nucleolin	NM_005381	augmentée
6	Interleukin enhancer binding factor 3	NM_004516	augmentée
7	cDNA FLJ30781 fis, clone FEBRA2000874, correspondant au gène codant l'adenylate cyclase 1 (brain)	AK055343 / NM_021116	diminuée
8	TIF1 beta zinc finger protein	X97548	augmentée
9	Likely ortholog of mouse tumor differentially expressed 1 TDE1L	NM_020755	diminuée
10	DKFZp434D1112_s1 434 (synonym: htes3) cDNA clone DKFZp434D1112 3'	AL039831	diminuée
11	N-MYC proto-oncogene	NM_005378	augmentée
12	Small nuclear ribonucleoprotein D2 polypeptide 16.5kDa	NM_004597	augmentée
13	DNA replication licensing factor MCM2	NM_004526	augmentée
14	RuvB-like DNA helicase TIP49b	NM_006666	augmentée
15	Immediate early protein ETR101	NM_004907	augmentée
16	RNA binding protein S1, serine-rich domain	NM_006711	augmentée
17	Ornithine decarboxylase 1	NM_002539	augmentée
18	Activity-related cytoskeleton-asso. protein (KIAA0278)	NM_015193	augmentée
19	Secretogranin II (chromogranin C)	NM_003469	diminuée
20	Structure specific recognition protein 1	NM_003146	augmentée
21	Collagen type VI, alpha 3 chain	NM_004369	diminuée
22	Small nuclear ribonucleoprotein polypeptides B and B1	NM_003091	augmentée
23	Acidic nuclear phosphoprotein 32 family, member B	NM_006401	augmentée
24	Non-POU domain containing, octamer-binding	NM_007363	augmentée
25	Peripheral myelin protein 22	NM_000304	diminuée
26	Small nuclear ribonucleoprotein polypeptide E	NM_003094	augmentée
27	Putative L-type neutral amino acid transporter (KIAA0436)	AB007896	diminuée
28	Fibrillarin	NM_001436	augmentée
29	Tripartite motif-containing 2	NM_015271	diminuée



30	DNA replication licensing factor MCM6	NM_005915	augmentée
31	Polypyrimidine tract binding protein 1	NM_002819	augmentée
32	Small nuclear ribonucleoprotein polypeptide A	NM_004596	augmentée
33	Creatine kinase, brain	NM_001823	augmentée
34	Erythrocyte membrane protein band 4.1 like3	NM_012307	diminuée
35	Hypothetical protein MGC3077	NM_024051	augmentée
36	Tissue alpha-L-fucosidase 1	NM_000147	diminuée
37	Secreted protein acidic and rich in cysteine	NM_003118	diminuée

Ces résultats ont également été validés par l'utilisation d'une autre technique de biologie moléculaire dans laquelle l'analyse de l'expression de gènes tel que présentés dans le tableau 2 a été réalisée par RT-PCR.

- 5 Pour cela, une réaction de reverse transcription (RT) a été réalisée à partir d'1 µg d'ARN total tel qu'obtenu précédemment (kit Amersham, First strand cDNA synthesis kit). La transcription reverse a été effectuée pendant 1 h à 37°C. Chaque solution de cDNA a été diluée 6 fois avant la réalisation de la PCR.

- L'expression des ARNm de gènes du tableau 2 (Peripheral myelin protein ou PMP22 (SEQ ID N°25) ; Insulin-like growth factor binding protein ou IGFBP7 (SEQ ID N°3 ; SPARC (SEQ ID N°37); EPB41L3 (SEQ ID N°34)) a ensuite été analysé par PCR (polymerase chain reaction) et l'utilisation d'amorces d'amplification spécifiques (amplification du gène PMP22 : brin sens : 5'-AGGGAGGAAG GGAAAACAGA-3' (SEQ ID N°38); brin antisens : 5'-TTAAGGCTCA ACACGAGGCT-3' (SEQ ID N°39) ; gène IGFBP7 : brin sens : 5'-CTTGAGCTGT GAGGTCATCG-3' (SEQ ID N°40); brin antisens : 5'-TATAGCTCGG CACCTTCACC-3' (SEQ ID N°41); gène SPARC : brin sens : 5'-CTGCCTGCCA CTGAGGGTTCC-3' (SEQ ID N°42) ; brin antisens : 5'-TCCAGGCAGA ACAACAAACC ATCC-3' (SEQ ID N°43) ; gène EPB41L3 : brin sens : 5'-ACCACCACCA CTACCCACAT-3' (SEQ ID N°44) ; brin antisens : 5'-TGGTTTTTCCT AACGGTTTGC-3' (SEQ ID N°45); gène beta actine : brin sens : 5'-TGTTGGCGTA CAGGTCTTTG C-3' (SEQ ID N°46); brin antisens : 5'-GCTACGAGCT GCCTGACGG-3' (SEQ ID N°47). L'expression du gène codant la β-actine a été utilisée comme contrôle. Trente cycles de PCR sont ensuite été réalisés en présence des différentes amorces d'amplification (0,2µM); de dNTPs (0,15mM,

Euromedex) et d'enzyme polymérase (Taq Polymerase ; 0,027U/ $\mu$ l; Perkin Elmer) (dénaturation 30'' à 94 °C, hybridation 1' à 60 °C ; polymérisation 1' à 72 °C).

Les résultats obtenus sont présentés dans le tableau 3 ci dessous, qui montre la corrélation existant entre les résultats obtenues par l'utilisation d'une biopuce et ceux obtenues par RT-PCR.

	Patients BP		Patients MP		p
	Résultats biopuce	Résultats RT PCR	Résultats biopuce	Résultats RT PCR	
SEQ ID N°37 : SPARC	3498	2,2	709	0,87	0,04
SEQ ID N°3 IGFBP7	4112	3,7	691	1,87	0,003
SEQ ID N° 34 EPB41L3	6041	4,2	986	1,5	0,001
SEQ ID N°25 PMP22	9051	3,8	2282	2,75	0,003

**Tableau 3**

Les résultats de RT-PCR, obtenus à partir de 15 patients BP et 8 patients MP, sont exprimés par le rapport de quantification relative entre les ARNm du gène cible et les ARNm du gène  $\beta$ -actine qui servait de contrôle. Les résultats sont exprimés par la moyenne des rapports obtenus pour chacun des groupes de patients. La corrélation des résultats obtenus d'une part avec la biopuce et d'autre part avec la technique en RT-PCR a été établie grâce au test de corrélation du Tau-B de Kendall. Les patients MP présentaient un niveau d'expression diminuée pour les gènes SPARC , IGFBP7 , EPB41L3, et PMP22, confirmant les résultats présentés dans le tableau 2.

L'expression des ARNm des gènes de SEQ ID N°37: SPARC ; SEQ ID N°2: UBE2C ; SEQ ID N°3: IGFBP7 ; SEQ ID N°8: TRIM28 ; SEQ ID N°22: SNRBP; SEQ ID N°25: PMP22; SEQ ID N°29: TRIM2 ; SEQ ID N°34: EPB41L3 ; SEQ ID N°7 : clone FEBRA2000874, a également été analysée par RT-PCR quantitative.

Les ADNc nécessaires à l'analyse de chacun de ces gènes cibles ont été obtenus à partir d'un microgramme d'ARN totaux (first-strand DNA synthesis kit, Amersham France). Après une dilution de 6 fois, 2,5 µl d'ADNc ont été utilisés en PCR en temps réel, en présence d'un couple d'amorce (300 nM) spécifique de chaque gène cible (cf tableau ci dessous) et du tampon SYBR-Green Master Mix.

SEQ ID AMPLIFIÉE	Amorce sens	Amorce antisens
SEQ ID N°37: SPARC	SEQ ID N°48 : 5'-CACATTAGGC TGTTGGTTCA AACT-3'	SEQ ID N°49 : 5'-CAGGATGCGC TGACCACTT-3'
SEQ ID N°2: UBE2C	SEQ ID N°50 : 5'-TCCTCAGCC CTGCTATCA-3'	SEQ ID N°51 : 5'-TTCAGGATGT CCAGGCATAT GT-3'
SEQ ID N°3: IGFBP7	SEQ ID N°52 : 5'-TGTCCTCATC TGGAACAAGG-3'	SEQ ID N°53 : 5'-GGCAGGAGTT CTGTCCTTTG-3'
SEQ ID N°7: clone FEBRA2000874	SEQ ID N°54 : 5'-TTTACATCCA GAGGCACGAC-3'	SEQ ID N°55 : 5'-CACGATGTCA GCAAACAGG-3'
SEQ ID N°8: TRIM28	SEQ ID N°56 : 5'-CAGGAAGGCT ATGGCTTTGG-3'	SEQ ID N°57 : 5'-CCGTTTCACA CCTGACACAT G-3'
SEQ ID N°22: SNRBP	SEQ ID N°58 : 5'-GCTGGACCGG AAGTAGGTTT CT-3'	SEQ ID N°59 : 5'-GCCGCTACCG GAAATGC-3'
SEQ ID N°25: PMP22	SEQ ID N°60 : 5'-GAOCCAGTGC ATCCAACAG A-3'	SEQ ID N°61 : 5'-GTGTGCGCGT AAAGCTTCAC-3'
SEQ ID N°29: TRIM2	SEQ ID N°62 : 5'-CAGTAACAAC CAATGTGTGC AG-3'	SEQ ID N°63 : 5'-TGCCAAAACG ACTTTTGAAC-3'
SEQ ID N°34: EPB41L3	SEQ ID N°64 : 5'-GTGGACCCT GCTAAGGAAA-3'	SEQ ID N°65 : 5'-CAGATAGTTG GGCAGGGTCT-3'
Gène de ménage HPRT1	SEQ ID N°66 : 5'-CACTGGCAAA ACAATGCAGA CT-3'	SEQ ID N°67 : 5'-CGACCTTGAC CATCTTTGGA TT-3'

Le volume total de réaction était de 15 µl. L'amplification en PCR a été réalisée en microplaques 96 puits, par l'utilisation du système ABI Prism 7000 Sequence Detection (Applied BioSystem USA). Le gène de référence HPRT1 ainsi que le gène cible étaient

analysés simultanément. Après 10 min de dénaturation à 95°C, l'amplification a été réalisée selon les conditions suivantes : 40 cycles de 15 secondes à 95°C suivie de 1 minutes à 60°C. Les expériences ont été réalisées en duplicat. La quantification a été réalisée par l'utilisation de la méthode des courbes standards et l'utilisation de la méthode comparative CT telle que recommandé par le fabricant. Les courbes standards ont été obtenues à partir de dilution d'ADNc de lignées cellulaires de neuroblastome, et réalisées pour chaque PCR. L'expression du gène cible a été déterminée par l'utilisation de ces courbes standards. L'expression relative de chaque gène cible a été définie par comparaison avec l'expression du gène de référence. Les tests de corrélation de Pearson et Spearman ont été utilisés pour calculer la corrélation entre les résultats obtenus sur puce, et les résultats obtenus par RT-PCR.

Les résultats sont présentés dans le tableau ci dessous.

	Test de Spearman	Test de Pearson
SEQ ID N°37: SPARC	p=0,0008	p=0,0163
SEQ ID N°2: UBE2C	p<0,0001	p<0,0001
SEQ ID N°3: IGFBP7	p=0,004	p=0,0977
SEQ ID N°7	p<0,0001	p= 0,001
SEQ ID N°8: TRIM28	p=0,0428	p=0,0058
SEQ ID N°22: SNRBP	p= 0,0664	p= 0,503
SEQ ID N°25: PMP22	p=0,021	p=0,0079
SEQ ID N°29: TRIM2	p=0,2983	p=0,232
SEQ ID N°34: EPB41L3	p=0,0123	p=0,0606

Ces résultats montraient une bonne corrélation ( $p < 0,05$ ) des résultats obtenus sur puce et par RT PCR, concernant en particulier les gènes de SEQ ID N°2, 7, 8, et 25, suggérant que ces 4 gènes sont particulièrement pertinents pour le pronostic du neuroblastome.

Les inventeurs ont également étudié l'expression simultanée des 37 gènes du tableau 2 pour obtenir un profil d'expression. Les résultats sont présentés dans la figure 1. On observe sur ce dendrogramme deux groupes ayant deux profils d'expression différents : un premier groupe permettant de classer les patients de bon pronostic (BP) et un deuxième groupe permettant de classer les patients de mauvais pronostic (MP).

Dans l'objectif de valider le pouvoir de discrimination du profil d'expression de ces 37 gènes, 6 tumeurs supplémentaires de patients «test » ont été analysées sans connaissance préalable de leur pronostic, et classées comme étant de bon pronostic «BP-test » ou de mauvais pronostic «MP-test » en fonction de l'analyse de leur profil d'expression. Leur bon classement a été vérifié ensuite selon leurs propriétés cliniques : tous les échantillons « tests » analysés en aveugle par l'analyse de l'expression de 37 gènes avaient été correctement classés dans le groupe de patients de mauvais pronostic « test-MP » ou dans le groupe de patients de bon pronostic « test-BP ». Ceci confirme que l'analyse de l'expression de ces 37 gènes est un bon outil pour le pronostic du neuroblastome.

A titre indicatif, l'oncogène N-MYC a également été utilisée comme outil de pronostic. L'utilisation de ce gène mettait en évidence 5 patients de mauvais pronostic (MP). Toutefois, 3 patients étaient également de mauvais pronostic alors qu'aucune augmentation de l'expression de l'oncogène N-MYC n'ait été observée, suggérant que l'analyse unique de ce gène n'est pas suffisant pour le pronostic du neuroblastome.

Les inventeurs ont également défini des panels de gènes plus restreint permettant également de discriminer les patients de bon et de mauvais pronostic.

Un premier panel comportait 19 gènes qui sont présentés dans le tableau 4. Les résultats sont exprimés par le ratio obtenu entre l'expression moyenne du gène chez des patients MP et l'expression du gène chez des patients BP (ratio MP / BP).

**Tableau 4 - liste des 19 gènes exprimés différemment dans les neuroblastomes de patients BP et MP**

SEQ ID N° .	Description de la séquence	N° Genbank	Ratio MP/BP
1	Flap structure-specific endonuclease 1 (FEN1)	NM_004111	2,7
2	Ubiquitin-conjugating enzyme E2C (UBE2C)	NM_007019	2,9
3	Insulin-like growth factor binding protein 7 (IGFBP7)	NM_001553	0,2
7	cDNA FLJ30781 fis, clone FEBRA2000874, correspondant au gène codant l'adenylate cyclase 1 (brain)	AK055343 / NM_021116	0,3
8	TIF1beta zinc finger protein	X97548	1,8
9	Likely ortholog of mouse tumor differentially expressed 1, like (TDE1L)	NM_020755	0,5
10	DKFZp434D1112_s1 434 (synonym: htes3) cDNA clone DKFZp434D1112 3'	AL039831	0,4
14	RuvB-like 2 (E. coli)(RUVBL2)	NM_006666	2,1
16	RNA binding protein S1, serine-rich domain (RNPS1), transcript variant 1	NM_006711	1,6
20	Structure specific recognition protein 1 (SSRP1)	NM_003146	2,0
21	Collagen, type VI, alpha 3 (COL6A3), transcript variant 1	NM_004369	0,2
22	Small nuclear ribonucleoprotein polypeptides B and B1 (SNRPB)	NM_003091	1,7
25	Peripheral myelin protein 22 (PMP22), transcript variant 1	NM_000304	0,3
27	KIAA0436 mRNA, partial cds	AB007896	0,5
29	Tripartite motif-containing 2 (TRIM2)	NM_015271	0,4
31	Polypyrimidine tract binding protein 1 (PTBP1), transcript variant 1	NM_002819	2,1
34	Erythrocyte membrane protein band 4.1-like 3	NM_012307	0,2
36	Fucosidase, alpha-L- 1, tissue (FUCA1)	NM_000147	0,2
37	Secreted protein, acidic, cysteine-rich (osteonectin) (SPARC)	NM_003118	0,2

Les inventeurs ont également étudié l'expression simultanée de ces 19 gènes pour obtenir un profil d'expression. Les résultats sont présentés dans la figure 2. On observe sur ce dendogramme deux groupes ayant deux profils d'expression différents : un premier groupe permettant de classer les patients de bon pronostic (BP) et un deuxième groupe permettant de classer les patients de mauvais pronostic (MP).

Dans l'objectif de valider le pouvoir de discrimination de ces 19 gènes, 6 tumeurs de patients « test » ont été analysées sans connaissance préalable de leur pronostic. Ainsi, les six tumeurs « MP-test » et « BP-test » présentées dans la figure 3 sont des tumeurs qui ont

été analysés « en aveugle ». Leur bon classement a été vérifié selon leur propriété clinique : tous les patients « MP-test » classés en fonction de leur profil d'expression comme patients de mauvais pronostic s'avéraient être des patients de mauvais pronostic et tous les patients « BP-test » classés en fonction de leur profil d'expression comme patients de bon pronostic s'avéraient être des patients de bon pronostic.

D'une façon comparable, un deuxième panel comportait 16 gènes tels que présentés dans le tableau 5.

**Tableau 5 - liste des 16 gènes exprimés différemment dans les neuroblastomes de patients BP et MP**

SEQ ID N°	Description de la séquence	N° Genbank
1	Flap structure-specific endonuclease 1 (FEN1)	NM_004111
2	Ubiquitin-conjugating enzyme E2C (UBE2C)	NM_007019
3	Insulin-like growth factor binding protein 7 (IGFBP7)	NM_001553
7	cDNA FLJ30781 fis, clone FEBRA2000874, correspondant au gène codant l'adenylate cyclase 1 (brain)	AK055343 / NM_021116
8	TIF1beta zinc finger protein	X97548
9	Likely ortholog of mouse tumor differentially expressed 1, like (TDE1L)	NM_020755
10	DKFZp434D1112_s1 434 (synonym: htes3) cDNA clone DKFZp434D1112 3'	AL039831
20	Structure specific recognition protein 1 (SSRP1)	NM_003146
21	Collagen, type VI, alpha 3 (COL6A3), transcript variant 1	NM_004369
22	Small nuclear ribonucleoprotein polypeptides B and B1 (SNRPB)	NM_003091
25	Peripheral myelin protein 22 (PMP22), transcript variant 1	NM_000304
29	Tripartite motif-containing 2 (TRIM2)	NM_015271
31	Polypyrimidine tract binding protein 1 (PTBP1), transcript variant 1	NM_002819
34	Erythrocyte membrane protein band 4.1-like 3 (EPB41L3)	NM_012307
36	Fucosidase, alpha-L-1, tissue (FUCA1)	NM_000147
37	Secreted protein, acidic, cysteine-rich (osteonectin) (SPARC)	NM_003118

Les inventeurs ont également étudié l'expression simultanée de ces 16 gènes pour obtenir un profil d'expression. Les résultats sont présentés dans la figure 3. On observe sur ce dendrogramme deux groupes ayant deux profils d'expression différents : un premier groupe

permettant de classer les patients de bon pronostic (BP) et un deuxième groupe permettant de classer les patients de mauvais pronostic (MP).

Un troisième panel comportait 12 gènes tels que présentés dans le tableau 6.

5 **Tableau 6 - liste des 12 gènes exprimés différemment dans les neuroblastomes de patients BP et MP**

SEQ ID N°	Description de la séquence	N° Genbank
2	Ubiquitin-conjugating enzyme E2C (UBE2C)	NM_007019
3	Insulin-like growth factor binding protein 7 (IGFBP7)	NM_001553
7	cDNA FLJ30781 fis, clone FEBRA2000874, correspondant au gène codant l'adenylate cyclase 1 (brain)	AK055343 / NM_021116
8	TIF1beta zinc finger protein	X97548
10	DKFZp434D1112_s1 434 (synonym: htes3) cDNA clone DKFZp434D1112 3'	AL039831
20	Structure specific recognition protein 1 (SSRP1)	NM_003146
22	Small nuclear ribonucleoprotein polypeptides B and B1 (SNRPB)	NM_003091
25	Peripheral myelin protein 22 (PMP22), transcript variant 1	NM_000304
29	Tripartite motif-containing 2 (TRIM2)	NM_015271
31	Polypyrimidine tract binding protein 1 (PTBP1), transcript variant 1	NM_002819
34	Erythrocyte membrane protein band 4.1-like 3 (EPB41L3)	NM_012307
37	Secreted protein, acidic, cysteine-rich (osteonectin) (SPARC)	NM_003118

Les inventeurs ont également étudié l'expression simultanée de ces 12 gènes pour obtenir un profil d'expression. Les résultats sont présentés dans la figure 4. On observe sur ce dendrogramme deux groupes ayant deux profils d'expression différents : un premier groupe permettant de classer les patients de bon pronostic (BP) et un deuxième groupe permettant de classer les patients de mauvais pronostic (MP).

Un quatrième panel comportait 9 gènes tels que présentés dans le tableau 7.

15 **Tableau 7 - liste des 9 gènes exprimés différemment dans les neuroblastomes de patients BP et MP**

SEQ ID N°	Description de la séquence	N° Genbank
2	Ubiquitin-conjugating enzyme E2C (UBE2C)	NM_007019



3	Insulin-like growth factor binding protein 7 (IGFBP7)	NM_001553
7	cDNA FLJ30781 fis, clone FEBRA2000874, correspondant au gène codant l'adenylate cyclase 1 (brain)	AK055343 / NM_021116
8	TIF1beta zinc finger protein	X97548
10	DKFZp434D1112_s1 434 (synonym: htes3) cDNA clone DKFZp434D1112 3'	AL039831
22	Small nuclear ribonucleoprotein polypeptides B and B1	NM_003091
25	Peripheral myelin protein 22 (PMP22), transcript variant 1	NM_000304
29	Tripartite motif-containing 2 (TRIM2)	NM_015271
34	Erythrocyte membrane protein band 4.1-like 3 (EPB41L3)	NM_012307

Les inventeurs ont également étudié l'expression simultanée de ces 9 gènes pour obtenir un profil d'expression. Les résultats sont présentés dans la figure 5. On observe sur ce dendrogramme deux groupes ayant deux profils d'expression différents : un premier groupe permettant de classer les patients de bon pronostic (BP) et un deuxième groupe permettant

5 de classer les patients de mauvais pronostic (MP).

Le pouvoir de discrimination de tous ces panels de gènes a été validé avec des tumeurs « tests » tels que décrit précédemment : tous les patients « MP-test » classés en fonction de leur profil d'expression comme patients de mauvais pronostic s'avéraient être des patients

10 de mauvais pronostic et tous les patients « BP-test » classés en fonction de leur profil d'expression comme patients de bon pronostic s'avéraient être des patients de bon pronostic.

Ces résultats démontrent que le pronostic d'un neuroblastome peut être déterminé par

15 l'analyse de l'expression de tout ou partie des 37 gènes de séquence SEQ ID N°1 à 37. En particulier, l'analyse de l'expression des 9 gènes de SEQ ID N°2 ; 3 ; 7 ; 8 ; 10 ; 22 ; 25 ; 29 ; et 34 permet de discriminer très efficacement les patients de bon et de mauvais pronostic.

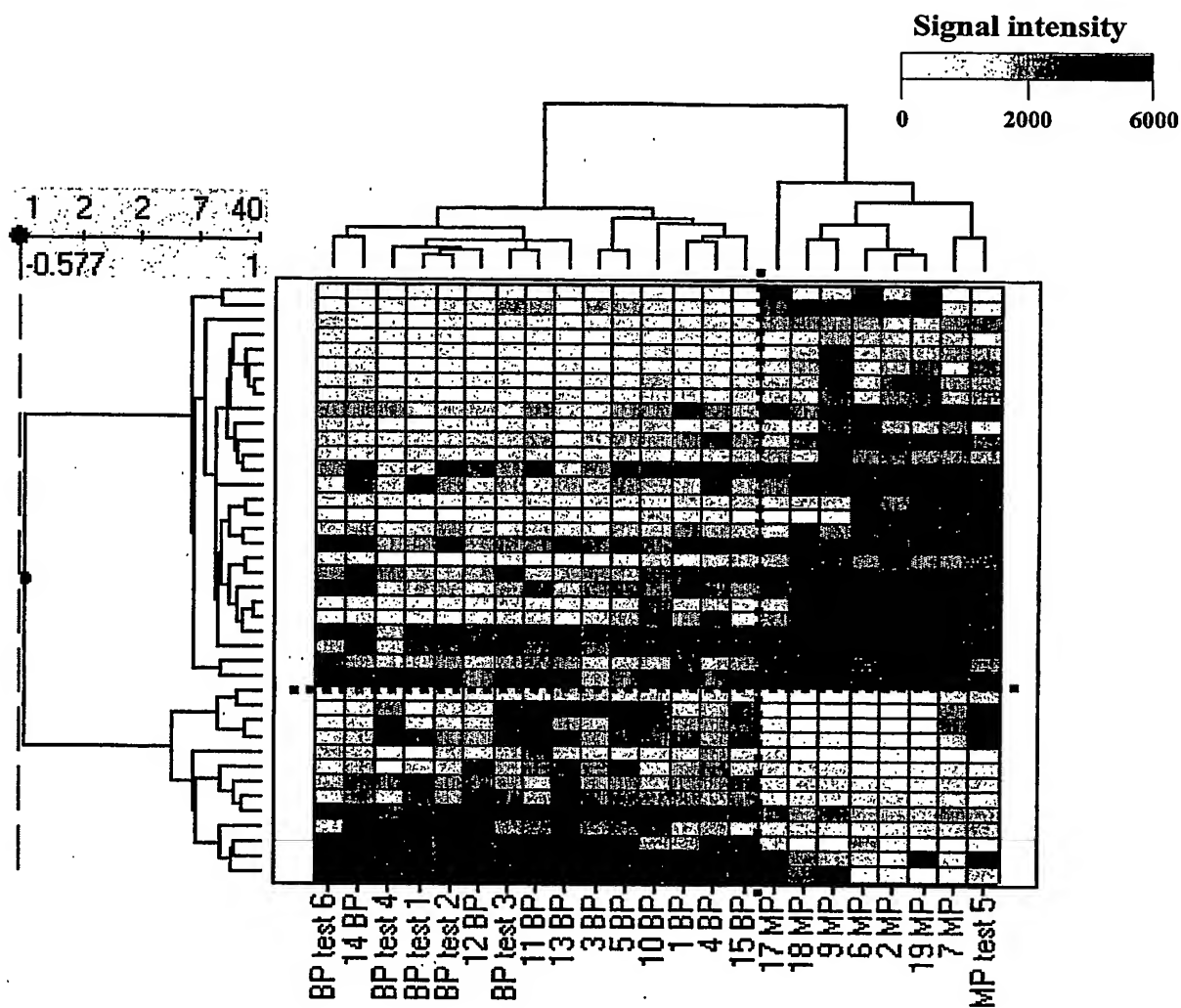
## REVENDICATIONS

1. Procédé pour le pronostic du neuroblastome chez un patient atteint du neuroblastome caractérisé en ce qu'il comprend les étapes suivantes :
  - 5 a. on extrait du matériel biologique d'un échantillon biologique prélevé chez le patient,
  - b. on met en contact le matériel biologique avec au moins un réactif spécifique choisi parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant l'une quelconque des SEQ ID N° 1 à 37, étant entendu que  
10 lorsque que le gène cible présente une séquence nucléique ayant l'une des SEQ ID N°11, 17 ou 37, on met en contact le matériel biologique avec au moins deux réactifs spécifiques choisis parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant l'une quelconque des SEQ ID N° 1 à 37,
  - 15 c. on détermine l'expression d'au moins un desdits gènes cibles, étant entendu que lorsque que le gène cible présente une séquence nucléique ayant l'une des SEQ ID N°11, 17 ou 37, on détermine l'expression d'au moins deux desdit gènes cibles.
- 20 2. Procédé pour le pronostic du neuroblastome selon la revendication 1 caractérisé en ce que l'échantillon biologique prélevé chez le patient est un échantillon tissulaire.
3. Procédé selon la revendication 1 ou 2 caractérisé en ce que le matériel biologique extrait lors de l'étape a) comprend des acides nucléiques.
- 25 4. Procédé selon la revendication 3 caractérisé en ce que le au moins un réactif spécifique de l'étape b) comprend au moins une sonde d'hybridation

5. Procédé selon la revendication 4 caractérisé en ce que la au moins une sonde d'hybridation est immobilisée sur un support.
6. Procédé selon la revendication 5 caractérisé en ce que le support est une biopuce.
- 5
7. Procédé selon l'une quelconque des revendications 1 à 6 caractérisé en ce que lors de l'étape b) on met en contact le matériel biologique avec au moins 37 réactifs spécifiques choisi parmi les réactifs spécifiques des gènes cibles présentant une séquence nucléique ayant l'une quelconque des SEQ ID N° 1 à 37 et on
- 10 détermine, lors de l'étape c, l'expression d'au moins 37 desdits gènes cibles.
8. Procédé selon l'une quelconque des revendications 1 à 6 caractérisé en ce que lors de l'étape b) on met en contact le matériel biologique avec au moins 19 réactifs spécifiques choisi parmi les réactifs spécifiques des gènes cibles présentant une
- 15 séquence nucléique ayant la SEQ ID N°1 ; SEQ ID N°2; SEQ ID N°3; SEQ ID N°7; SEQ ID N°8; SEQ ID N°9; SEQ ID N°10; SEQ ID N°14; SEQ ID N°16; SEQ ID N°20; SEQ ID N°21; SEQ ID N°22; SEQ ID N°25; SEQ ID N°27; SEQ ID N°29; SEQ ID N°31; SEQ ID N°34; SEQ ID N°36 ou SEQ ID N°37 et on détermine, lors de l'étape c) l'expression d'au moins 19 desdits gènes cibles.
- 20
9. Procédé selon l'une quelconque des revendications 1 à 6 caractérisé en ce que lors de l'étape b) on met en contact le matériel biologique avec au moins 9 réactifs spécifiques choisi parmi les réactifs spécifiques des gènes cibles présentant une
- 25 séquence nucléique ayant la SEQ ID N°2; SEQ ID N°3; SEQ ID N°7; SEQ ID N°8; SEQ ID N°10; SEQ ID N°22; SEQ ID N°25; SEQ ID N°29; SEQ ID N°34; ou SEQ ID N°37 et on détermine, lors de l'étape c) l'expression d'au moins 9 desdits gènes cibles.

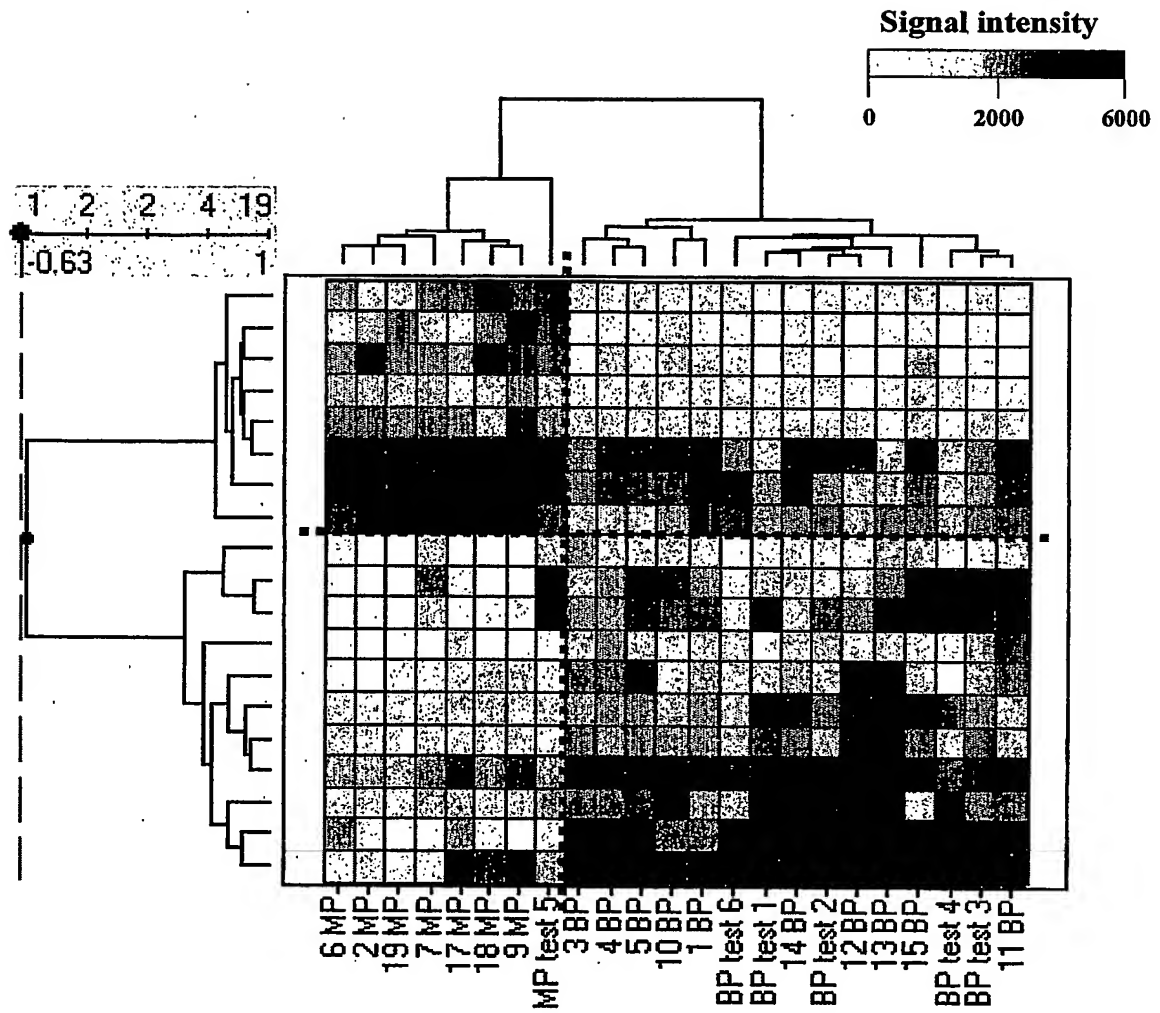
1/5

Figure 1



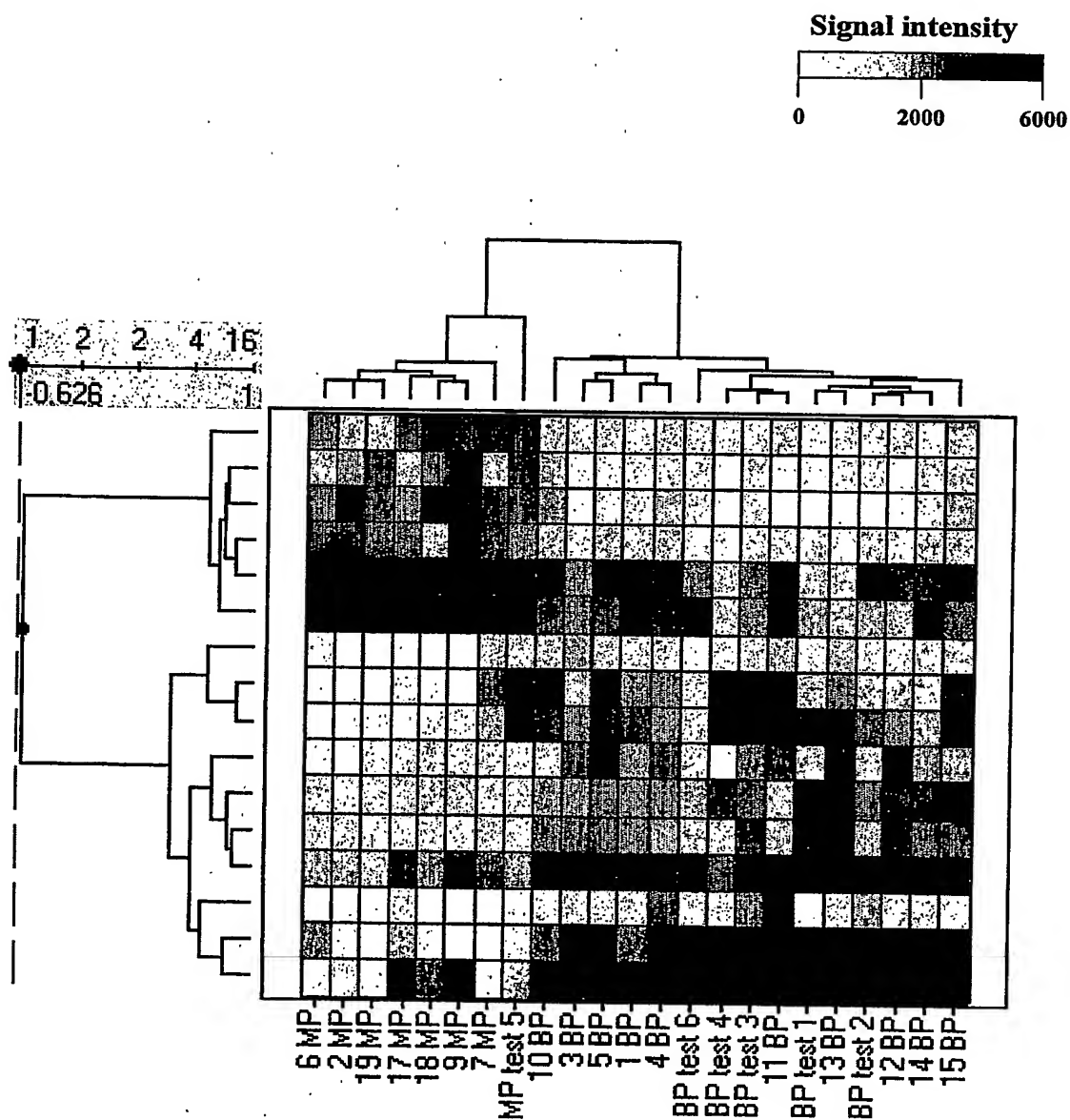
2/5

Figure 2



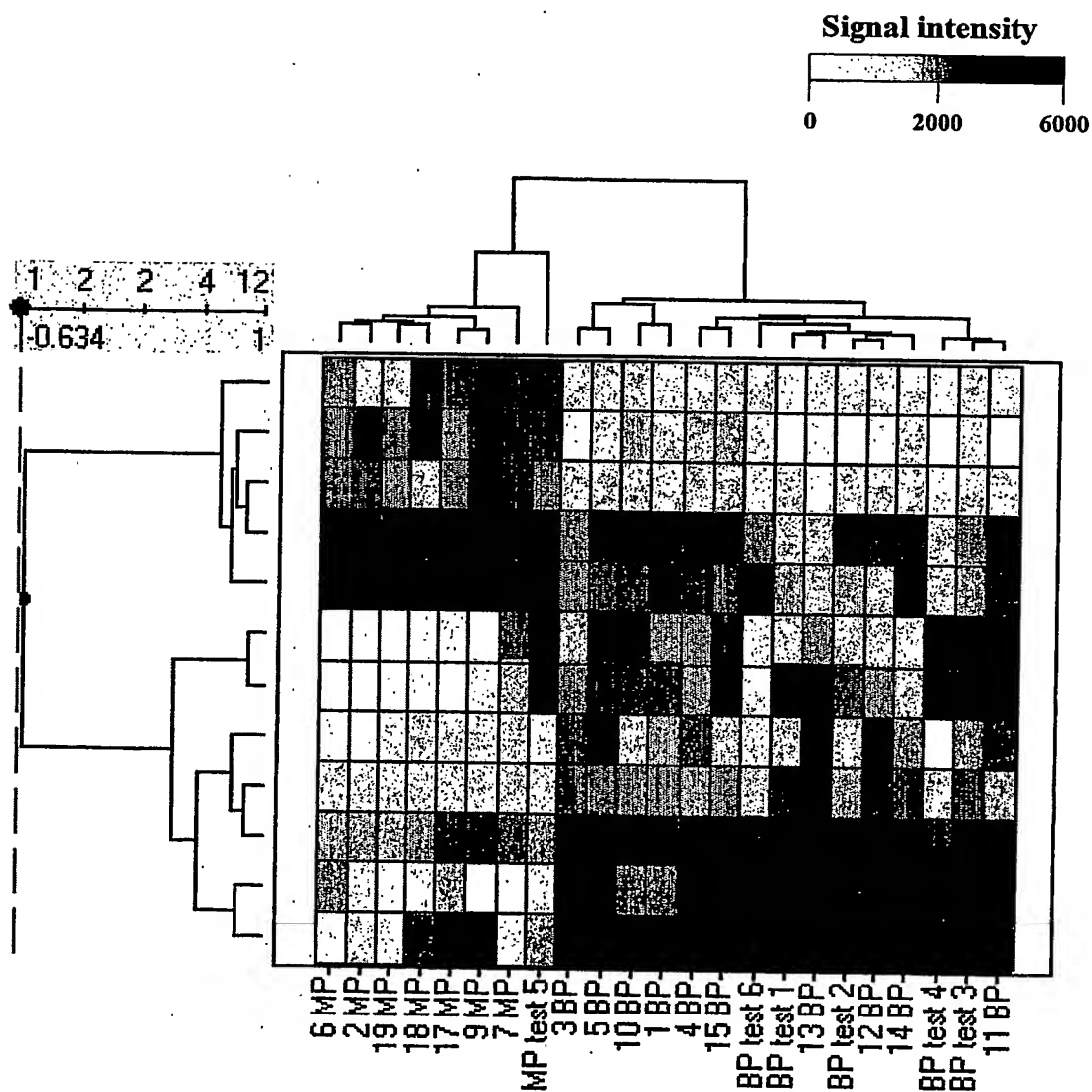
3/5

Figure 3



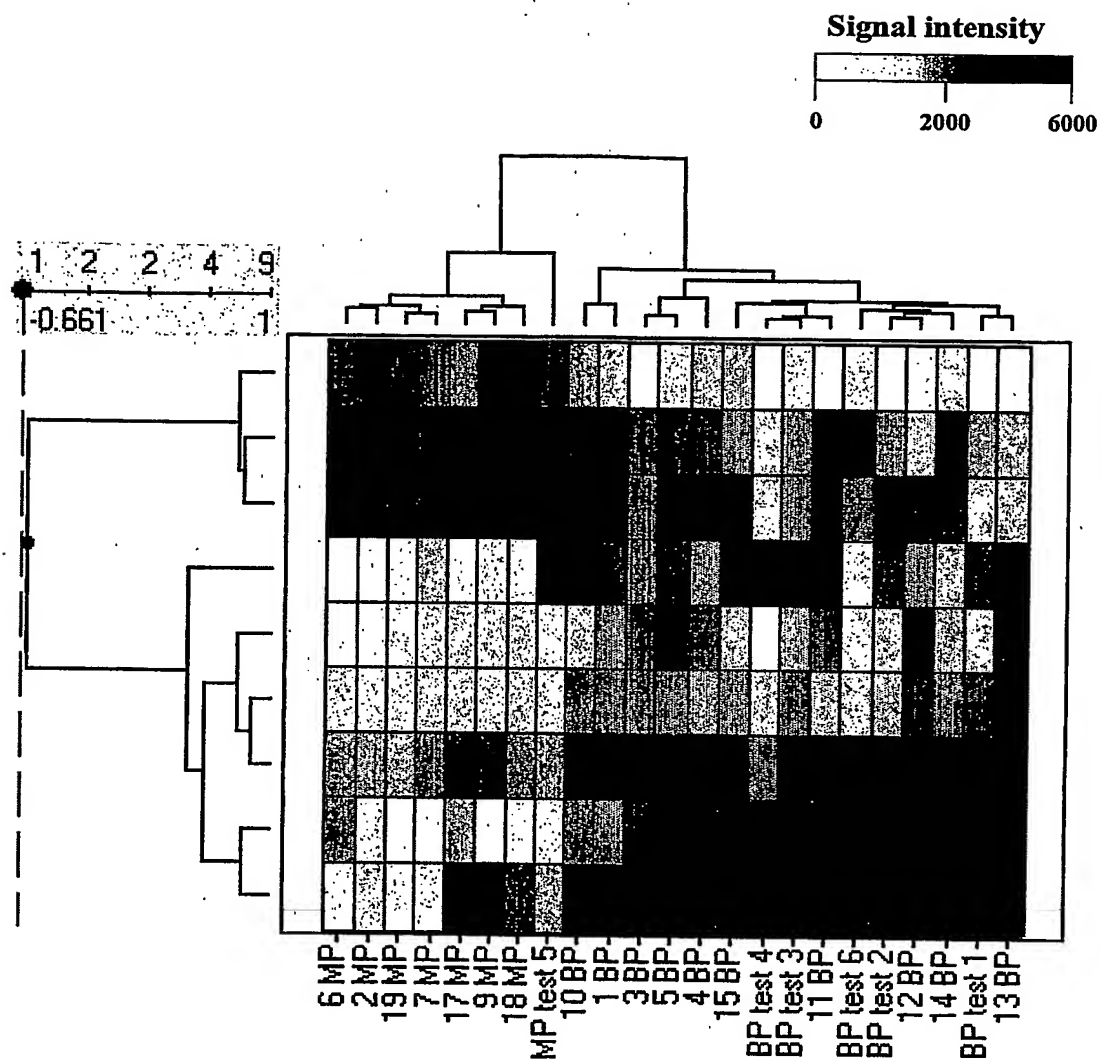
4/5

Figure 4



5/5

Figure 5





## SEQUENCE LISTING

&lt;110&gt; BIOMERIEUX SA / CENTRE LEON BERARD

&lt;120&gt; Procédé pour le pronostic du neuroblastome

&lt;130&gt; Unknown

&lt;160&gt; 67

&lt;170&gt; PatentIn version 3.1

&lt;210&gt; 1

&lt;211&gt; 2265

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

```

<400> 1
agtcctgcga tttcgggtgt agagggagca ggggcctgcg gggacctggt gtgggtggag      60
tggggacaag cggtggagaa gggtagcca gggcgctga gagactctgt tctccctgga      120
gggactgggt gccatgagag cagccgtctg aggggacgca gcctgcacta cgcgccccaa      180
gaggctgtgc gtggcgagca ggtcacgtga cgggagcgcg ggctttggaa ggcggctgaa      240
cgtcaggcca cccgccgcta agctgagaag ggagagcgag cttaggaccg cctgcccggg      300
gcaaccccga accaagcttt agccgccgag gccgcgtgtc ccaaaggcca gtcatccctc      360
ctctgtgttg ccatgggaat tcaaggcctg gccaaactaa ttgctgatgt ggcccccagt      420
gccatccggg agaatgacat caagagctac tttggccgta aggtggccat tgatgcctct      480
atgagcattt atcagttcct gattgctgtt cgccagggtg gggatgtgct gcagaatgag      540
gagggtgaga ccaccagcca cctgatgggc atgttctacc gcaccattcg catgatggag      600
aacggcatca agcccgtgta tgtctttgat ggcaagccgc cacagctcaa gtcaggcgag      660
ctggccaaac gcagtgagcg gcgggctgag gcagagaagc agctgcagca ggctcaggct      720
gctggggccg agcaggaggt ggaaaaattc actaagcggc tggatgaagg cactaagcag      780
cacaatgatg agtgcaaaca tctgctgagc ctcatgggca tcccttatct tgatgcaccc      840
agtgaggcag aggccagctg tgctgccctg gtgaaggctg gcaaagtcta tgctgcggct      900
accgaggaca tggactgcct caccttcggc agccctgtgc taatgcgaca cctgactgcc      960
agtgaagcca aaaagctgcc aatccaggaa ttccacctga gccggattct gcaggagctg     1020
ggcctgaacc aggaacagtt tgtggatctg tgcatcctgc taggcagtga ctactgtgag     1080

```

```

agtatccggg gtattgggcc caagcgggct gtggacctca tccagaagca caagagcatc 1140
gaggagatcg tgcggcgact tgaccccaac aagtaccctg tgccagaaaa ttggctccac 1200
aaggaggctc accagctctt cttggaacct gaggtgctgg acccagagtc tgtggagctg 1260
aagtggagcg agccaaatga agaagagctg atcaagttca tgtgtggtga aaagcagttc 1320
tctgaggagc gaatccgcag tggggtcaag aggctgagta agagccgcca aggcagcacc 1380
cagggccgcc tggatgattt cttcaaggty accggctcac tctcttcagc taagcgcaag 1440
gagccagaac ccaagggatc cactaagaag aaggcaaaga ctggggcagc agggaagttt 1500
aaaaggggaa aataaatgtg tttcccccatt atacctcctt caccacagaa tatttgccgt 1560
cttgtaccct taagagctac agctagagaa accttcacgg ggtggagaga ggattctaag 1620
gcttttctag cgtgaccctt ttcagtagtg ctatgccctt ttttacttga tcttaatggc 1680
aagaaggcca cagaggctact tttccttttt tagctcagga aaatatgtca ggctcaaacc 1740
acttctcagg cagtttaatg gacactaagt ccattgttac atgaaagtga tagatagcaa 1800
caagtttttg agaagagaga gggagataaa agggggagac aaaagatgta cagaaatgat 1860
ttcctggctg gccaactggg ggccagtggg aggtgatggg ggacctagac tgtgcttttc 1920
tgtcttggtc agccttgacc caccttgaga gagagccacc aggaaggcgc atcttagcag 1980
atgggaggaa ctgctgagag aagatgggca gaaagctgga gcccttgag ttggctgtgt 2040
ctgtgtttgt gactgattac tggctgtgtc ttgggtgggc agaaactcga acttgctatg 2100
taatttgtgt ctagttattc agaggagtaa gatggtgatg ttcacctggc aatcagctga 2160
gttgagactt tggaataaga cactggtttt catgcgctgt ttttgtttta aagttatgaa 2220
gaaaaaagtc aataaaattc taaaagtaaa aaaaaaaaaa aaaaa 2265

```

<210> 2

<211> 783

<212> DNA

<213> Homo sapiens

<400> 2

```

ggcacgagcg agttcctgtc tctctgcaa cgccgcccgg atggcttccc aaaaccgcga 60
cccagccgcc actagcgtcg ccgcccggc taaaggagct gagccgagcg ggggcgcccgc 120
ccggggtccg gtgggcaaaa ggctacagca ggagctgatg accctcatga tgtctggcga 180
taaagggatt tctgccttcc ctgaatcaga caaccttttc aaatgggtag ggaccatcca 240
tggagcagct ggaacagtat atgaagacct gaggtataag ctctcgctag agttccccag 300
tggctaccct tacaatgcmc ccacagtga gttcctcacg ccctgctatc accccaacgt 360
ggacacccag ggtaacatat gcctggacat cctgaaggaa aagtgggtctg ccctgtatga 420
tgtcaggacc attctgctct ccatccagag cttcttagga gaacccaaca ttgatagtcc 480
cttgaacaca catgctgccg agctctggaa aaaccccaca gcttttaaga agtacctgca 540

```

```

agaaacctac tcaaagcagg tcaccagcca ggagccctga cccaggctgc ccagcctgtc 600
cttgtgtcgt ctttttaatt tttccttaga tggctgtgcc tttttgtgat ttctgtatag 660
gactctttat cttgagctgt ggtatttttg ttttgTTTTT gtcttttaaa ttaagcctcg 720
gttgagccct tgtatatata ataatgcat ttttgcctt ttttaaaaaa aaaaaaaaaa 780
aaa 783

```

<210> 3

<211> 1124

<212> DNA

<213> Homo sapiens

```

<400> 3
gccgctgcca ccgcaccccg ccatggagcg gccgtcgtg cgcgccctgc tcctcggcgc 60
cgctgggctg ctgctcctgc tcctgcccct ctctcttcc tcctcttcgg acacctgcgg 120
cccctgcgag ccggcctcct gcccgcccct gccccgctg ggctgcctgc tgggcgagac 180
ccgcgacgcg tgcggctgct gccctatgtg cgcccgcggc gagggcgagc cgtgcggggg 240
tggcggcgcc ggcagggggg actgcgcgcc gggcatggag tgcgtgaaga gccgcaagag 300
gcggaagggg aaagccgggg cagcagccgg cgggccgggt gtaagcggcg tgtgcgtgtg 360
caagagccgc taccgggtgt gcggcagcga cggcaccacc taccgagcg gctgccagct 420
gcgcgccgcc agccagaggg ccgagagccg cggggagaag gccatcacc aggtcagcaa 480
gggcacctgc gagcaaggtc cttccatagt gacgcccccc aaggacatct ggaatgtcac 540
tggtgcccag gtgtacttga gctgtgaggt catcggaatc ccgacacctg tcctcatctg 600
gaacaaggta aaaaggggtc actatggagt tcaaaggaca gaactcctgc ctggtgaccg 660
ggacaacctg gccattcaga cccgggggtg ccagaaaaag catgaagtaa ctggctgggt 720
gctggtatct cctctaagta aggaagatgc tggagaatat gagtgccatg catccaattc 780
ccaaggacag gcttcagcat cagcaaaaat tacagtgggt gatgccttac atgaaatacc 840
agtgaaaaaa ggtgaagggt ccgagctata aacctccaga atattattag tctgcatggt 900
taaaagtagt catggataac tacattacct gttcttgctt aataagtttc ttttaatcca 960
atccactaac actttagtta tattcactgg ttttacacag agaaatacaa aataaagatc 1020
acacatcaag actatctaca aaaatttatt atatatttac agaagaaaag catgcatatc 1080
attaaacaaa taaaatactt tttatcacia aaaaaaaaaa aaaa 1124

```

<210> 4

<211> 5084

<212> DNA

<213> Homo sapiens

<400> 4  
 agcaccacgg cagcaggagg tttcggctaa gttggaggta ctggccacga ctgcatgccc 60  
 gcgcccacca ggtgatacct ccgccggtga cccaggggct ctgcgacaca aggagtctgc 120  
 atgtctaagt gctagacatg ctcagctttg tggatacgcg gactttgttg ctgcttgag 180  
 taaccttatg cctagcaaca tgccaatctt tacaagagga aactgtaaga aaggggccag 240  
 ccggagatag aggaccacgt ggagaaagg gtcaccagg ccccccaggc agagatgggtg 300  
 aagatgggtcc cacaggccct cctgggtccac ctgggtcctcc tggccccctt ggtctcggtg 360  
 ggaactttgc tgctcagtat gatggaaaag gagttggact tggccctgga ccaatgggct 420  
 taatgggacc tagaggccca cctgggtgcag ctggagcccc aggccctcaa ggtttccaag 480  
 gacctgctgg tgagcctggt gaacctggtc aaactgggtcc tgcaaggctc cgtgggtccag 540  
 ctggccctcc tggcaaggct ggtgaagatg gtcaccctgg aaaaccgga cgacctgggtg 600  
 agagaggagt tgttggaaca cagggtgctc gtggtttccc tggaaactcct ggacttcctg 660  
 gcttcaaagg cattagggga cacaatggtc tggatggatt gaaggagacag cccggtgctc 720  
 ctggtgtgaa ggggtgaacct ggtgcccctg gtgaaaatgg aactccaggc caaacaggag 780  
 cccgtgggct tcctgggtgag agaggacgtg ttggtgcccc tggcccagct ggtgcccgtg 840  
 gcagtgatgg aagtgtgggt cccgtgggtc ctgctgggtc cattgggtct gctggccctc 900  
 caggcttccc aggtgcccct ggcccaagg gtgaaattgg agctgttggg aacgctgggtc 960  
 ctgctgggtc cgccgggtccc cgtgggtgaag tgggtcttcc aggcctctcc ggccccgttg 1020  
 gacctcctgg taatcctgga gcaaacggcc ttactgggtc caagggtgct gctggccttc 1080  
 ccggcggttc tgggggtccc ggcctccctg gaccccgcg tattcctggc cctgttgggtg 1140  
 ctgccgggtc tactgggtgc agaggacttg ttggtgagcc tgggtccagct ggctccaaag 1200  
 gagagagcgg taacaagggt gagcccggtc ctgctggggc ccaagggtcct cctgggtcca 1260  
 gtggtgaaga aggaaagaga ggcctaata gggaaagctg atctgccggc cctccaggac 1320  
 ctctggggt gagaggtagt cctggttctc gtggtcttcc tggagctgat ggcagagctg 1380  
 gcgtcatggg ccctcctggt agtcgtgggt caagtggccc tgctggagtc cgaggaccta 1440  
 atggagatgc tggtcgccct ggggagcctg gtctcatggg acccagaggc cttcctgggt 1500  
 cccctggaaa tatcggtccc gctggaaaag aaggctcctg cggcctccct ggcacgacg 1560  
 gcaggcctgg cccaattggc ccagctggag caagaggaga gcctggcaac attggattcc 1620  
 ctggacccaa aggccccact ggtgatcctg gcaaaaacgg tgataaaggc catgctgggtc 1680  
 ttgctgggtc tcgggggtgct ccaggctcct atggaaacaa tgggtgctcag ggacctcctg 1740  
 gaccacaggg tgttcaagggt ggaaaagggt aacagggtcc cgctgggtcct ccaggcttcc 1800  
 agggctctgc tggccccctc ggtcccgtg gtgaagttgg caaacaggga gaaaggggtc 1860  
 tccatggtga gtttgggtct cctggtcctg ctggtccaag aggggaacgc ggtccccag 1920  
 gtgagagtgg tgctgccggt cctactggtc ctattggaag ccgaggtcct tctggacccc 1980  
 cagggcctga tggaaacaag ggtgaacctg gtgtgggttg tgctgtgggc actgctgggtc 2040

catctggtcc	tagtggactc	ccaggagaga	ggggtgctgc	tggcatacct	ggaggcaagg	2100
gagaaaaggg	tgaacctggt	ctcagagggtg	aaattggtaa	ccctggcaga	gatggtgctc	2160
gtggtgctca	tgggtgctgta	ggtgccccctg	gtcctgctgg	agccacaggt	gaccggggcg	2220
aagctggggc	tgctggctct	gctggtcctg	ctggtcctcg	gggaagccct	ggtgaacgtg	2280
gcgaggctcg	tcctgctggc	cccaacggat	ttgtggtcc	ggctggtgct	gctggtcaac	2340
cgggtgctaa	aggagaaaga	ggagccaaag	ggcctaaggg	tgaaaacggt	gttggtggtc	2400
ccacaggccc	cggtggagct	gctggcccag	ctggtccaaa	tgggtccccc	ggtcctgctg	2460
gaagtctggt	tgatggaggc	ccccctggta	tgactggttt	ccctggtgct	gctggacgga	2520
ctggtccccc	aggacctct	ggtatttctg	gccctcctgg	tccccctggt	cctgctggga	2580
aagaagggct	tcgtggctct	cgtggtgacc	aagggtccagt	tggccgaact	ggagaagtag	2640
gtgcagttgg	tccccctggc	ttcgctggtg	agaaggggtcc	ctctggagag	gctggtactg	2700
ctggacctcc	tggcactcca	ggtcctcagg	gtcttcttgg	tgctcctggt	attctgggtc	2760
tccctggctc	gagagggtgaa	cgtggtctac	ctggtgttgc	tgggtgctgtg	ggtgaacctg	2820
gtcctcttgg	cattgccggc	cctcctgggg	cccgtggtcc	tcctggtgct	gtgggtagtc	2880
ctggagtgca	cggtgctcct	ggtgaagctg	gtcgtgatgg	caaccctggg	aacgatggtc	2940
ccccaggctc	cgatggtcaa	ccgggacaca	agggagagcg	cggttaccct	ggcaatattg	3000
gtcccgttgg	tgctgcaggt	gcacctggtc	ctcatggccc	cgtgggtcct	gctggcaaac	3060
atggaaaccg	tggtgaaact	ggtccttctg	gtcctgttgg	tcctgctggt	gctgttggcc	3120
caagagggtcc	tagtggccca	caaggcattc	gtggcgataa	gggagagccc	ggtgaaaagg	3180
ggcccagagg	tcttcctggc	ttaaagggac	acaatggatt	gcaagggtctg	cctggtatcg	3240
ctggtcacca	tggtgatcaa	ggtgctcctg	gctccgtggg	tcctgctggt	cctagggggc	3300
ctgctggtcc	ttctggccct	gctggaaaag	atggtcgcac	tggacatcct	ggtacggttg	3360
gacctgctgg	cattcgaggc	cctcaggggtc	accaaggccc	tgctggcccc	cctggtcccc	3420
ctggccctcc	tggacctcca	ggtgtaagcg	gtggtggtta	tgactttggt	tacgatggag	3480
acttctacag	ggctgaccag	cctcgctcag	caccttctct	cagacccaag	gactatgaag	3540
ttgatgctac	tctgaagtct	ctcaacaacc	agattgagac	ccttcttact	cctgaaggct	3600
ctagaaagaa	cccagctcgc	acatgccgtg	acttgagact	cagccacca	gagtggagca	3660
gtggttacta	ctggattgac	cctaaccaag	gatgcactat	ggatgctatc	aaagtatact	3720
gtgatttctc	tactggcgaa	acctgtatcc	gggcccaccc	tgaaaacatc	ccagccaaga	3780
actggtatag	gagctccaag	gacaagaaac	acgtctggct	aggagaaact	atcaatgctg	3840
gcagccagtt	tgaatataat	gtagaaggag	tgacttccaa	ggaaatggct	acccaacttg	3900
ccttcatgct	cctgctggcc	aactatgcct	ctcagaacat	cacctaccac	tgcaagaaca	3960
gcattgcata	catggatgag	gagactggca	acctgaaaaa	ggctgtcatt	ctacagggtc	4020
ctaattgatgt	tgaacttggt	gctgagggca	acagcagggt	cacttacact	gttctttag	4080
atggctgctc	taaaaagaca	aatgaatggg	gaaagacaat	cattgaatac	aaaacaata	4140

```

agccatcacg cctgcccttc cttgatattg cacctttgga catcggtggt gctgaccatg 4200
aattctttgt ggacattggc ccagtctgtt tcaaataaat gaactcaatc taaattaaaa 4260
aagaaagaaa ttgaaaaaa ctttctcttt gccatttctt cttcttcttt tttaactgaa 4320
agctgaatcc ttccatttct tctgcacatc tacttgctta aattgtgggc aaaagagaaa 4380
aagaaggatt gatcagagca ttgtgcaata cagtttcatt aactccttcc cccgctcccc 4440
caaaaatttg aatttttttt tcaacactct tacacctgtt atggaaaatg tcaacctttg 4500
taagaaaacc aaaataaaaa ttgaaaaata aaaaccataa acatttgcac cacttggtggc 4560
ttttgaatat ctccacaga gggaagttaa aaacccaaac ttccaaagggt ttaaactacc 4620
tcaaaacact ttcccatgag tgtgatccac attgttaggt gctgacctag acagagatga 4680
actgagggcc ttgttttgtt ttgttcataa tacaagggtg ctaattaata gtatttcaga 4740
tactgaaga atgttgatgg tgctagaaga atttgagaag aaatactcct gtattgagtt 4800
gtatcgtgtg gtgtattttt taaaaaattt gatttagcat tcatattttc catcttattc 4860
ccaattaaaa gtatgcagat tatttgccca agttgtcct cttcttcaga ttcagcattt 4920
gttctttgcc agtctcattt tcatcttctt ccatggttcc acagaagctt tgtttcttgg 4980
gcaagcagaa aaattaaatt gtacctattt tgtatatgtg agatgtttaa ataaattgtg 5040
aaaaaatga aataaagcat gtttggtttt caaaagaac atat 5084

```

<210> 5

<211> 2518

<212> DNA

<213> Homo sapiens

<400> 5

```

cttcgggtgt acgtgctccg ggatcttcag caccgcggc cgccatcgcc gtcgcttggc 60
ttcttctgga ctcatctgcg ccacttgctc gttcacact ccgccgccat catggtgaag 120
ctcgcgaagg caggtaaaaa tcaagggtgac cccaagaaaa tggctcctcc tccaaaggag 180
gtagaagaag atagtgaaga tgaggaaatg tcagaagatg aagaagatga tagcagtgga 240
gaagaggctg tcatacctca gaagaaaggc aagaaggctg ctgcaacctc agcaaagaag 300
gtggtcgttt cccaacaaa aaaggttgca gttgccacac cagccaagaa agcagctgtc 360
actccaggca aaaaggcagc agcaacacct gccagaaga cagttacacc agccaaagca 420
gttaccacac ctggcaagaa gggagccaca ccaggcaaag cattggtagc aactcctggt 480
aagaagggtg ctgccatccc agccaagggg gcaagaatg gcaagaatgc caagaaggaa 540
gacagtgatg aagaggagga tgatgacagt gaggaggatg aggaggatga cgaggacgag 600
gatgaggatg aagatgaaat tgaaccagca gcgatgaaag cagcagctgc tgcccctgcc 660
tcagaggatg aggacgatga ggatgacgaa gatgatgagg atgacgatga cgatgaggaa 720
gatgactctg aagaagaagc tatggagact acaccagcca aaggaaagaa agctgcaaaa 780

```

gttggttcctg	tgaaagccaa	gaacgtggct	gaggatgaag	atgaagaaga	ggatgatgag	840
gacgaggatg	acgacgacga	cgaagatgat	gaagatgatg	atgatgaaga	tgatgaggag	900
gaggaagaag	aggaggagga	agagcctgtc	aaagaagcac	ctggaaaacg	aaagaaggaa	960
atggccaaac	agaaagcagc	tcctgaagcc	aagaaacaga	aagtggagg	cacagaaccg	1020
actacggctt	tcaatctctt	tggttgaaac	ctaaacttta	acaaatctgc	tcctgaatta	1080
aaaactggta	tcagcgatgt	ttttgctaaa	aatgatcttg	ctggttgga	gtcagaatt	1140
ggtatgacta	ggaaatttgg	ttatgtggat	tttgaatctg	ctgaagacct	ggagaaagcg	1200
ttggaactca	ctggtttgaa	agtctttggc	aatgaaatta	aactagagaa	accaaaggga	1260
aaagacagta	agaaagagcg	agatgcgaga	acacttttgg	ctaaaaatct	cccttacaaa	1320
gtcactcagg	atgaattgaa	agaagtgttt	gaagatgctg	cggagatcag	attagtcagc	1380
aaggatggga	aaagtaaagg	gattgcttat	attgaattta	agacagaagc	tgatgcagag	1440
aaaacctttg	aagaaaagca	gggaacagag	atcgatgggc	gatctatttc	cctgtactat	1500
actggagaga	aaggtcaaaa	tcaagactat	agaggtggaa	agaatagcac	ttggagtggg	1560
gaatcaaaaa	ctctggtttt	aagcaacctc	tcctacagtg	caacagaaga	aactcttcag	1620
gaagtatttg	agaaagcaac	ttttatcaaa	gtaccccaga	accaaattgg	caaatctaaa	1680
gggtatgcat	ttatagagtt	tgcttcattc	gaagacgcta	aagaagcttt	aaattcctgt	1740
aataaaaggg	aaattgaggg	cagagcaatc	aggctggagt	tgcaaggacc	caggggatca	1800
cctaagcca	gaagccagcc	atccaaaact	ctgtttgtca	aaggcctgtc	tgaggatacc	1860
actgaagaga	cattaaagga	gtcatttgac	ggctccgttc	gggcaaggat	agttactgac	1920
cgggaaactg	ggctctccaa	agggtttggg	tttgtagact	tcaacagtga	ggaggatgcc	1980
aaggaggcca	tggaagacgg	tgaaattgat	ggaaataaag	ttaccttgga	ctgggccaaa	2040
cctaagggtg	aaggtggcct	cgggggtcgt	ggtggaggca	gaggcggcct	tggaggacga	2100
ggtggtggta	gaggaggccg	aggaggattt	ggtggcagag	gccggggagg	ctttggaggg	2160
cgaggaggct	tccgaggagg	cagaggagga	ggaggtgacc	acaagccaca	aggaaagaag	2220
acgaagtttg	aatagcttct	gtccctctgc	tttccctttt	ccatttgaaa	gaaaggactc	2280
tggggttttt	actgttacct	gatcaatgac	agagccttct	gaggacattc	caagacagta	2340
tacagtcctg	tggtctcctt	ggaaatccgt	ctagttaaca	tttcaagggc	aataccgtgt	2400
tggttttgac	tggaatttca	tataaacttt	ttaaagagtt	gagtgataga	gctaaccctt	2460
atctgtaagt	tttgaattta	tattgtttca	tcccatgtac	aaaaccattt	tttcctac	2518

&lt;210&gt; 6

&lt;211&gt; 3677

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

```

<400> 6
cgccctgcccc cccgccccgct cgcccccggt ccggactcct cctcctcctc ttctcgccat    60
tgcagttgga cccagcagcc cggcgcgcac cgctggcctt ttgggggcag accccggcgg    120
gctgtggcag gagggcggcg gcggcggtg cggtcgaaga aggggacgcc gacaagagtt    180
gaagtattga taacaccaag gaactctatc acaatttgaa aagataagca aaagtttgat    240
ttccagacac tacagaagaa gtaaaaatgc gtccaatgcg aatttttgat aatgatgacc    300
gccatgtgat ggcaaagcat tcttccgttt atccaacaca agaggagctg gaggcagtcc    360
agaacatggt gtcccacacg gagcgggcgc tcaaagctgt gtccgactgg atagacgagc    420
aggaaaaggg tagcagcagc caggcagagt ccgataacat ggatgtgccc ccagaggacg    480
acagtaaaga aggggctggg gaacagaaga cggagcacat gaccagaacc ctgcggggag    540
tgatgcgggt gggcctggtg gcaaagggcc tcctactcaa gggggacttg gatctggagc    600
tggtgctgct gtgtaaggag aagcccaca ccgcccctc ggacaagggt gccgacaacc    660
tggccatcca gcttgctgct gtaacagaag acaagtacga aatactgcaa tctgtcgacg    720
atgctgcatg tgtgataaaa aacacaaaag agcctccatt gtccctgacc atccacctga    780
catccccctg tgtcagagaa gaaatggaga aagtattagc tggagaaacg ctatcagtca    840
acgaccccc ggacgttctg gacaggcaga aatgccttgc tgccttggcg tccctccgac    900
acgccaagtg gttccaggcc agagccaacg ggctgaagtc ttgtgtcatt gtgatccggg    960
tcttgaggga cctgtgact cgctgcccc cctggggtcc cctccgaggc tggcctctcg   1020
agctcctgtg tgagaaatcc attggcacgg ccaacagacc gatgggtgct ggcgaggccc   1080
tgcggagagt gctggagtgc ctggcgctcg gcatcgtgat gccagatggt tctggcattt   1140
atgacccttg tgaaaaagaa gccactgatg ctattgggca tctagacaga cagcaacggg   1200
aagatatcac acagagtgcg cagcacgcac tgcggctcgc tgccttcggc cagctccata   1260
aagtcctagg catggaccct ctgccttcca agatgcccc gaaaccaaag aatgaaaacc   1320
cagtggacta caccgttcag atcccaccaa gcaccaccta tgccattacg cccatgaaac   1380
gccaatgga ggaggacggg gaggagaagt cggccagcaa aaagaagaag aagattcaga   1440
agaaagagga gaaggcagag cccccccagg ctatgaatgc cctgatgcgg ttgaaccagc   1500
tgaagccagg gctgcagtac aagctgggtg cccagactgg gcccgctccat gccccatct   1560
ttaccatgtc tgtggagggt gatggcaatt cattcgaggc ctctgggccc tccaaaaaga   1620
cggccaagct gcacgtggcc gttaagggtg tacaggacat gggcttgccg acgggtgctg   1680
aaggcaggga ctcgagcaag ggggaggact cggctgagga gaccgaggcg aagccagcag   1740
tggtggcccc tgccccagt gtagaagctg tctccacccc tagtgcggcc tttccctcag   1800
atgccactgc cgagcagggg ccgatcctga caaagcacgg caagaacca gtcatggagc   1860
tgaacgagaa gaggcgtggg ctcaagtacg agctcatctc cgagaccggg ggcagccacg   1920
acaagcgctt cgtcatggag gtcgaagtgg atggacagaa gttccaagggt gctggttcca   1980
acaaaaaggt ggcgaaggcc tacgctgctc ttgctgccct agaaaagctt ttcctgaca   2040
cccctctcgc cttgatgcc aacaaaaaga agagagcccc agtaccgctc agagggggac   2100

```



cgaaatttgc tgctaagcca cataaccctg gcttcggcat gggaggcccc atgcacaacg 2160  
 aagtgcaccc accccccaac cttcgagggc ggggaagagg cgggagcatc cggggacgag 2220  
 ggcgcgggag aggatttggt ggcgccaacc atggaggcta catgaatgcc ggtgctgggt 2280  
 atggaagcta tgggtacgga ggcaactcgg cgacagcagg ctacagtacg tttttcacag 2340  
 actgctacgg ctatcatgat tttgggtcct cctagagcgt ctaaaagtat tgcacacaaa 2400  
 atcaactttt tactccaatt tcctccaact ccaaaaccca aagtgtccgt gctgtgtccc 2460  
 tgtgcttcac tgggtttctc aaccgtggct tttcacgca gcttgtctga aactcttagc 2520  
 ctgcagaatt taagacaatg gcagttttta tcgtgatttg cctttgaact tggcctatt 2580  
 gaagttcaca ataagtggaa aacaattttt tcagagaatg tttttttgtg cagaattgca 2640  
 cagaattcta gagacagcgt tgttcggcat caaggcaaaa gccacacctt gctttttatg 2700  
 gaaagcatta ctttttttaa agagacagac aatgacgcat tttaatctac ctttgtctta 2760  
 atttacagca ggttttgtat gaatttttaa ctttttaaca aactcccaaa tctggttgat 2820  
 gcctttgaca gtgatgaaaa cgatttcacc acatctgaat ccagagaaac cggctttttt 2880  
 tcttattgag agcatgttaa aacgttggga acatgtgggg aattgtatat tgcgctgaat 2940  
 taacttctcc cgcctcttgt aatgctctgg tgggttcttg tttgggaatg cgatattttg 3000  
 tggctgggtt agctagagag tgaactctca aaggatcaa aactgtgctt ccattattag 3060  
 tgcaagaaac agacaggcct taaggggtag atgacgtgaa attttgcaag tcttaattac 3120  
 agctgcagat gcatgggatt ctggattttt ttgttgcttt ttagtttaat gggactttta 3180  
 aagtaattga ggagaaagaa ccgtgatgtt ccctgtttct ccagtaaagg actggctttt 3240  
 gcttgggcag aggtggtgct gctgggtgtg cagctgccac agactccaaa ggcgtagaag 3300  
 tttgtgcaa cacacggagt cattctggct ctctgctgag gcccctgtt tctggcagg 3360  
 gccctccttg gaaactggtt ttggctctga tcagcgggtt tttttgcagc aaagcctgca 3420  
 tctgtgttga cttgcaagat tttgcgttta ttcaggcaaa aactgggtcaa aatgggtact 3480  
 acatgatttg ttcccagagg tttgaaacat tcagtgaac tttttaaaac tttgattgca 3540  
 tgatgtattt tttttttaga aagttattgt ttgagaataa tgtcttttta taccaggaaa 3600  
 atagttatcc tgaatgacgt tgaaaactcc ccctccctt tttttttttt taatcaatac 3660  
 atgtgaaagt aacaagc 3677

<210> 7

<211> 2901

<212> DNA

<213> Homo sapiens

<400> 7

ttgaaatcag gaaatcaggc cgggcgcagt ggctcatgcc tgtaagccca gcactttggg 60  
 aggcggaggc ggggtgatcc attgagggtc ggagctcaag accagcctgg tcaacatggt 120

gaaaccccgt ctctactaaa aatacaaaaa aaaaattagc tgggcgtgtt ggcgggagcc	180
tgtagtccca gctacacggg aggctaaggt gggagaattg cttgaaccg ggaggcggag	240
gttgacgtga gctgagattg caccattgca ctccagcctg ggcgacagag caagactctc	300
tcaaaaaaaaa aaaaaagaa agaaagaaat cagaaaatcg accacagtgg tagccacctg	360
gcctaagtgt gtgtttttgt acctgacagg ggtcactcat tttaggcaca actccttcat	420
tctttgtgaa attagttagt ttccttctac ccgtcaccag attcaatatg ttctattaat	480
acaccgataa ccacagggga agggcacttg tcgctctccc acctggttac cacagtctcc	540
atgggtcttt tgccgtgacc acaataaag gaaacactca tcactagtat ctaagtcggg	600
ctttacagta actatgcacc ttctgtgtgc ttcacctcac tctctacttc aaacagccca	660
tggagggagg tattattata ctctttatgt tgacagtga gaatctgagg cccagagagg	720
ttggggactt gagtaaagtc acacagccct gagaggcagg accagggttc cattcctgct	780
ctatccagtt ccaagccctt gtgttttcca ttatgtttag tgcctctttg ctaacagcaa	840
catctgcaag atttgtgttg gttttgatgg agaactctag ctcatccaca tgctagtgcc	900
caagtgtgtg agggggccacc tcagcaggtg ggttctgaat gcagccaagg ctgtccccgc	960
aatgggtgag actcgctcca actgcccggc ctgagagcag gtgcctaagt cctccctggc	1020
actggcaggc cttacctcac attgctaaat taaagcaatg caattcctct tgggtaagag	1080
gaattcctcc ttctttacta actgatcccc agcaaggaaa taaaatgtta ggctttaaaa	1140
atccctactt tgtcatatca gactatattc taaaactata tttagcgaa acctgtcatt	1200
gcgtctaatt tcaaatatac agaatctcct taagagctgt tgccttattt ttttgtaaag	1260
cctctctgac atcaaatggg gagaaatggg ggcacctca gacaccctga aactacacac	1320
catttcttcc ctgctcagct tctgctcagg agttctgtga gctatgggaa ggccattggt	1380
tgtatttgct acttttactt tcatcttcct ctgctgtaga gccatttaat gttattgtca	1440
tatgctgctg gtgaggtaaa ggtgggtccg ggtgccttcc caggggtag aggatgttca	1500
aagggccgat ttcagcagga gttcagaggg cttatgatgg atggtagag atttgacaac	1560
caccagagca catgtgctct gaccctctcc tgggcattgg ttcctgctgg taccgggcgg	1620
ttcagacctt caaatagggt gctttcaaaa gagctttcag gcacttattg agaattaatg	1680
tttaaacaga cataatagcc tagatgaact cccaagagat ctattaaatc ttgtgggctg	1740
aataaatatc tcgtgcagga ctgtgcaaca gtagcccaga gcatcctgcc tgtgggcac	1800
cacctccag gtgagggcag tgggaagctg gcccgcaggc agccagaact tgtttctcac	1860
ctcccaccag caacccccca cccaactctg ggccccaggc acacgaagca caagtctcag	1920
gggaccattc ccacattggg ggatcctgag ggagcccatc accgcctctt gcatacaact	1980
gtccactagg aggcacgccc agtgtgggag agatgtatgg tcttgcttc cacctgtaaa	2040
aactgcacat atgcaagcca tttgcactct ggaactgcat gccgtgaaaa ctctaatgg	2100
tgtggaactt agtttgaatt tgaaatcacg ccgcatgcac aaaggacag gccaggccc	2160
gacctcaggt catccgcccg ctggctgcag agcatccctg ggagccaagg cgaggcccg	2220

ggagcctgag ctttgtgtag ctcgagcttt gtgtagctcg tgcacttatt atgcaccacc 2280  
 tcccttcagt caccactcct cttcctccgc catcctcatt tatactgatt gcacaccccc 2340  
 cgctcaaaca acaatgtcct tattatgatg accatctcgt agtggtagat tccattccta 2400  
 ttttaaggtaa gcccaaagcc cacttttgga ttttctcgac tgtccgagaa aagttgtgta 2460  
 agcgctgagc ttcttctggg tttggctaga tagggttgtg tccctctatg gaatggagag 2520  
 tgatgtgggc aagggtgtca ttttctcgca caatacaact cactgaggat gcttctgtag 2580  
 aagtgagaaa cacgatgagt acattcagaa ttacaataac tcactctcac tgggtaactt 2640  
 ctcagtatag atttgtatga tcaatacggg tctattttta tgtcaactga aactgttagg 2700  
 gtaccttcca gtctttttca agattgttaa attgagacaa gtaattgaat aatttgcctt 2760  
 atttttatct taaaaaagt gaatggactg aaatgttaaa tgtgaatga catttcttaa 2820  
 ttgcaatttt tctactgagt gtttgcacta tactttctgg aatcttattt aacaaaaata 2880  
 aagggaataa attgcttgac t 2901

<210> 8

<211> 3056

<212> DNA

<213> Homo sapiens

<400> 8

gcggggcggg ccggcggcgg aggcggggcc gcggagccag gagtgactag cagcagttgg 60  
 ccgtgccgta gcagcgtccc gcgcgcggcg ggcagcggcc caggaggcgc gtggtgcggg 120  
 tttcggcggc ggctgaggaa gaagcgcggg cggcgccttc gggaggcgag caggcagcag 180  
 ttggccgtgc cgtagcagcg tcccgcgcg gcggggcagc ggcccaggag gcgcgtggcg 240  
 gcgctcggcc tcgcggcggc ggcggcggca gcggcccagc agttggcggc gagcgcgtct 300  
 gcgcctgcgc ggcgggcccc gcgcccctcc tccccctctg ggcgcccccg gcggcgtgtg 360  
 aatggcggcc tccgcggcgg cagcctcggc agcagcggcc tcggccgcct ctggcagccc 420  
 gggcccgggc gagggctccg ctggcggcga aaagcgcctc accgcccctt cggccgcagc 480  
 ctcggcctct gcctcagccg cggcgtcgtc gcccgcgggg ggcggcgccg aggcgtgga 540  
 gctgctggag cactgcggcg tgtgcagaga gcgcctgcga cccgagaggg agccccgcct 600  
 gctgccctgt ttgcactcgg cctgtagtgc ctgcttaggg cccgcggccc ccgccgccgc 660  
 caacagctcg ggggacggcg gggcggcggg cgacggcacc gtggtggact gtcccgtgtg 720  
 caagcaacag tgcttctcca aagacatcgt ggagaattat ttcattgcgtg atagtggcag 780  
 caaggctgcc accgacgccc aggatgcgaa ccagtgtgc actagctgtg aggataatgc 840  
 cccagccacc agctactgtg tggagtgtc ggagcctctg tgtgagacct gtgtagaggc 900  
 gcaccagcgg gtgaagtaca ccaaggacca tactgtgcgc tctactgggc cagccaagtc 960  
 tcgggatggt gaacgtactg tctattgcaa cgtacacaag catgaacccc ttgtgctggt 1020

ttgtgagagc	tgtgatactc	tcacctgccg	agactgccag	ctcaatgccc	acaaggacca	1080
ccagtaccag	ttcttagagg	atgcagttag	gaaccagcgc	aagctcctgg	cctcactggt	1140
gaagcgcctt	ggggacaaac	atgcaacatt	gcagaagagc	accaaggagg	ttcgcagctc	1200
aatccgccag	gtgtctgacg	tacagaagcg	tgtgcaagtg	gatgtcaaga	tggccatcct	1260
gcagatcatg	aaggagctga	ataagcgggg	ccgtgtgctg	gtcaatgatg	cccagaaggt	1320
gactgagggg	cagcaggagc	gcctggagcg	gcagcactgg	accatgacca	agatccagaa	1380
gcaccaggag	cacattctgc	gctttgcctc	ttgggctctg	gagagtga	acaacacagc	1440
ccttttgctt	tctaagaagt	tgatctactt	ccagctgcac	cgggccctca	agatgattgt	1500
ggatcccgtg	gagccacatg	gcgagatgaa	gtttcagtgg	gacctcaatg	cctggaccaa	1560
gagtgccgag	gcctttggca	agattgtggc	agagcgtcct	ggcactaact	caacaggccc	1620
tgcacccatg	gccccccaa	gagccccagg	gcccctgagc	aagcagggct	ctggcagcag	1680
ccagcccatg	gaggtgcagg	aaggctatgg	ctttgggtca	ggagatgata	cctactcaag	1740
tgcagagccc	catgtgtcag	gtgtgaaacg	gtcccgtcga	ggtgagggcg	aggtgagcgg	1800
ccttatgcgc	aagggtccac	gagtgcagct	tgaacgcctg	gacctggacc	tcacagctga	1860
cagccagcca	cccgtcttca	aggtcttccc	aggcagtacc	actgaggact	acaaccttat	1920
tgttattgaa	cgtggcgctg	ccgctgcagc	taccggccag	ccagggactg	cgctgcagg	1980
aacccctggt	gccccacccc	tggctggcat	ggccattgtc	aaggaggagg	agacggaggc	2040
tgccattgga	gccccctcta	ctgccactga	gggccctgag	accaaacctg	tgcttatggc	2100
tcttgcgagg	ggtcctgggt	ctgaggggtc	ccgcttgccc	tcacctagt	gcagcaccag	2160
ctcagggctg	gaggtgggtg	ctcctgaggg	tacctcagcc	ccaggtgggt	gccccgggaa	2220
cctggatgac	agtgccacca	tttgccgtgt	ctgccagaag	ccaggcgatc	tggttatgtg	2280
caaccagtgt	gagttttgtt	tccacctgga	ctgtcacctg	ccggccctgc	aggatgtacc	2340
aggggaggag	tggagctgct	cactctgcca	tgtgtccctt	gacctgaagg	aggaggatgg	2400
cagcctcagc	ctggatgggt	cagacagcac	tggcgtgggt	gccaagctct	caccagccaa	2460
ccagcggaaa	tgtgagcgtg	tactgctggc	cctattctgt	cacgaacctt	gcccggccct	2520
gcatcagctg	gctaccgact	ccaccttctc	cctggaccag	cccggtgcca	ccctggatct	2580
gacctgatc	cgtgcccgcc	tccaggagaa	gttgtcacct	ccctacagct	ccccacagga	2640
gtttgcccag	gatgtggggc	gcatgttcaa	gcaattcaac	aagttaactg	aggacaaggc	2700
agacgtgcag	tccatcatcg	gcctgcagcg	cttcttcgag	acgcgcatga	acgaggcctt	2760
cgggtgacac	aagttctctg	ctgtgctggg	ggagcccccg	ccgatgagcc	tgcttgggtg	2820
tggcctgagt	tcccaggagc	tgtctgggtg	ccctgggtgat	ggccccctgag	gctggagccc	2880
ccatggccag	cccagcctgg	ctctgttctc	tgtcctgtca	ccccatcccc	actcccctgg	2940
tggcctgact	cccactccct	ggtggcccca	tccccagtt	cctcacgata	tggtttttac	3000
ttctgtggat	ttaataaaaa	aaacttcacc	agttcaaaaa	aaaaaaaaaa	aaaaaa	3056

&lt;210&gt; 9

&lt;211&gt; 3149

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 9

```

agcggaaatct cggaaaggcg agaaagaagc tgtctccatc ttgtctgtat ccgctgctct    60
tgtgacgttg tggagatggg gagcgtcctg gggctgtgct ccatggcgag ctggatacca    120
tgtttggtg gaagtgtccc gtgtttgcta tgccgatgct gtcctagtgg aaacaactcc    180
actgtaacta gattgatcta tgcacttttc ttgcttggtg gagtatgtgt agcttggtga    240
atgttgatac caggaatgga agaacaactg aataagattc ctggattttg tgagaatgag    300
aaagggttg tcccttgtaa cattttgggt ggctataaag ctgtatatcg tttgtgcttt    360
ggtttggtc tgttctatct tcttctctct ttactaatga tcaaagtga gagtagcagt    420
gatcctagag ctgcagtgc caatggattt tggttcttta aatttgctgc agcaattgca    480
attattattg gggcattctt cattccagaa ggaactttta caactgtgtg gttttatgta    540
ggcatggcag gtgccttttg tttcatctc atacaactag tcttacttat tgattttgca    600
cattcatgga atgaatcgtg ggttgaaaa atggaagaag ggaactcgag atgttggtat    660
gcagccttgt tatcagctac agctctgaat tatctgctgt ctttagttgc tatcgtcctg    720
ttctttgtct actacactca tccagccagt tgttcagaaa acaaggcgtt catcagtgtc    780
aacatgctcc tctgcgttgg tgcttctgta atgtctatac tgccaaaaat ccaagaatca    840
caaccaagat ctggtttggt acagtcttca gtaattacag tctacacaat gtatttgaca    900
tggtcagcta tgaccaatga accagaaaca aattgcaacc caagtctact aagcataatt    960
ggctacaata caacaagcac tgtcccaaag gaagggcagt cagtccagtg gtggcatgct   1020
caaggaatta taggactaat tctctttttg ttgtgtgtat ttatttcag catccgtact   1080
tcaacaata gtcagggtta taaactgact ctaacaagtg atgaatctac attaatagaa   1140
gatggtggag ctagaagtga tggatcactg gaggatgggg acgatgttca ccgagctgta   1200
gataatgaaa gggatggtgt cacttacagt tattccttct ttcacttcat gcttttcctg   1260
gcttcacttt atatcatgat gacccttacc aactggtaca ggtatgaacc ctctcgtgag   1320
atgaaaagtc agtggacagc tgtctgggtg aaaatctctt ccagttggat tggcatcgtg   1380
ctgtatgttt ggacactcgt ggcaccactt gttcttacia atcgtgattt tgactgagtg   1440
agacttctag catgaaagtc ccactttgat tattgcttat ttgaaaacag tattcccaac   1500
ttttgtaaag ttgtgtatgt ttttgcttcc catgtaactt ctccagtgtt ctggcatgaa   1560
ttagatttta ctgcttgta ttttggtatt ttcttaccia gtgcattgat atgtgaagta   1620
gaatgaattg cagaggaaag ttttatgaat atggtgatga gttagtaaaa gtggccacta   1680
ttgggcttat tctctgctct atagtgtgta aatgaagagt gaaaacaaat ttgtttgact   1740
attttaaaat tatattagac cttagctgt ttagcaagc attaaagcaa atgtatggct   1800

```

```

gccttttaaa atatttgatg tgttgccctgg caggatactg caaagaacat ggttttatfff 1860
aaaatttata aacaagtcac tttaatgcc a gttgtctgaa aaatcttata aggtttttacc 1920
cttgatacgg aattttacaca ggtagggagt gtttagtgga caatagtgtg ggttatggat 1980
ggaggtgtcg gtactaaatt gaataacgag taaataatct tacttgggta gagatggcct 2040
ttgccaacaa agtgaactgt tttggttgtt ttaaaactcat gaagtatggg ttcagtggaa 2100
atgttttgaa ctctgaagga tttagacaag gttttgaaaa ggataatcat gggttagaag 2160
gaagtgtttg aaagtcactt tgaaagttag ttttgggcca gcacggtagc tcacccttgt 2220
aatcccagca ctttgggagg ctgaggtggg tagattactt gagcccagga attcaagacc 2280
agcctgggca acatggtgaa accctgtttc tataaaaaat aatctgggct ttgtagcata 2340
tgcctgtggt cccagctact gaggaggctg aggtgggagg attgcttgag cccaggaggc 2400
agaggttgca gtgagccaag gtcacgtcac tgcactctag cctgggcaac agagtaagac 2460
aaaaaaatat atatatattg aaaatcaaag gaggcaaaat tttgacaggg aaggaagtaa 2520
ctgcaaaaca ctaggcttta gtaggtactt atataaaatc tagtccagtt ctctcattta 2580
aaaaaatgaa gacactgaag tacagactta aatagctcag atagctaatt aggaaatttc 2640
aagttggcca ataatagcat tctctctgac atttaaaaaat aatttctatt caaaatacat 2700
gcataattga ttttacacct cattactggt ggataattta tgtgatgtgg attgctggtg 2760
tccagcatga ccataaaca ggtcagaaga atgatggaat gttttagaat aaactcctgc 2820
ttatagtata ctacacagtt caaaagatgt ttaaaatgct tttgtattta ctgccatgta 2880
attgaaatat atagattatt gtaacctttc aacctgaaaa tcaagcagta tgagagttta 2940
gttatttgta tgtgtcacta gtgtctaata aagcttttaa aatctacaat ttcttcttta 3000
aaaatattta ttaatgtgaa tggaatataa caattcagct taattcccca accttattct 3060
gtgtgtagac attgtattcc acaattttga atggctgtgt tttacctcta aataaatgaa 3120
ttcagagaaa gtgaaaaaaa aaaaaaaaaa 3149

```

&lt;210&gt; 10

&lt;211&gt; 580

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 10

```

cttttttttt ttttttttta aagtcttttag tatattttatt tgtataaaga gtaaacaag 60
tgcatataga gtggccacag gtttgacaca gagaccttg tgatgtaggc tatgaacaaa 120
tttaaatggc aacttcattg ctgccactga accaatcctg aatttgggct caacagggtga 180
aaagtaacaa tatcaaacga atactaaaca gcataacaaa aagattttca gactcttggg 240
cataaagacc gtaatcggtc acattgaatc aatgactaaa cttttttgat taccagcta 300
cctccaagca aactgaaaac tgtctagtgg atcctgaagt ccatagtgcc tctagccggg 360

```

tctttcaagt gttgcaccac aggggtgatga ttgatggtaa aaacagggat caacccttgt 420  
 agatcggtag taagtatgga aaccctctaa gaacagtga gcgtatgttg tattcagact 480  
 ggttgcatac agcattcaaa accagtgtg gaatagcttg ccccaaagt gtagagttat 540  
 aaaaggatat acattgacgt ttcttaaaag catgtgtaat 580

<210> 11

<211> 2467

<212> DNA

<213> Homo sapiens

<400> 11  
 ggcacgaggc tccggtgtgt ctgtcggttg cagtgttga ggtcggcgcc ggccccgcc 60  
 ttccgcgccc cccacgggaa ggaagcacc ccggtattaa aacgaacgg gcggaagaa 120  
 gccctcagtc gccggccggg aggcgagccg atgccgagct gctccacgt caccatgccg 180  
 ggcagatct gcaagaaccc agacctcgag ttgactcgc tacagccctg cttctacccg 240  
 gacgaagatg acttctactt cggcggcccc gactcgaccc cccgggggga ggacatctgg 300  
 aagaagtgtg agctgtgcc cagcggccc ctgtcgccca gccgtggctt cgcgagcac 360  
 agctccgagc ccccgagctg ggtcacggag atgtgtcttg agaacgagct gtggggcagc 420  
 ccggccgagg aggacgcgtt cggcctgggg ggactgggtg gcctcaccac caaccggtc 480  
 atcctccagg actgcatgtg gagcggcttc tccgcccgcg agaagctgga gcgcgccgtg 540  
 agcgagaagc tgcagcacgg ccgcgggccc ccaaccgccc gttccaccgc ccagtccccg 600  
 ggagccggcg ccgccagccc tgcgggtcgc gggcacggcg gggctgcggg agccggccgc 660  
 gccggggccg ccctgccgc cgagctcgcc caccggccg ccgagtgcgt ggatcccgcc 720  
 gtggtcttcc ctttcccggt gaacaagcgc gagccagcgc ccgtgcccgc agccccggcc 780  
 agtggcccgg cggcggggcc tgcgggtgcc tcggggggcg gtattgccgc cccagccggg 840  
 gccccggggg tcgcccctcc gcgccaggc ggccgccaga ccagcggcg cgaccacaag 900  
 gccctcagta cctccggaga ggacaccctg agcgattcag atgatgaaga tgatgaagag 960  
 gaagatgaag aggaagaaat cgacgtggtc actgtggaga agcggcggtt ctctccaac 1020  
 accaaggctg tcaccacatt caccatcact gtgcgtccca agaacgcagc cctgggtccc 1080  
 gggagggtc agtcagcga gctgatctc aaacgatgcc ttcccatcca ccagcagcac 1140  
 aactatgccg cccctctcc ctacgtggag agtgaggatg caccaccaca gaagaagata 1200  
 aagagcgagg cgtccccacg tccgctcaag agtgtatcc cccaaaggc taagagcttg 1260  
 agccccgaa actctgactc ggaggacagt gagcgtcga gaaaccacaa catcctggag 1320  
 cgccagcgcc gcaacgacct tcggtccagc tttctcacgc tcagggacca cgtgccggag 1380  
 ttggtaaaga atgagaaggc cgccaagggt gtcattttga aaaaggccac tgagtatgtc 1440  
 cactccctcc aggccgagga gcaccagctt ttgctggaaa aggaaaaatt gcaggcaaga 1500

```

cagcagcagt tgctaaagaa aattgaacac gctcggactt gctagacgct tctcaaaact 1560
ggacagtcac tgccactttg cacattttga tttttttttt aaacaaacat tgtgttgaca 1620
ttaagaatgt tggtttactt tcaaatcggt cccctgtcga gttcggctct ggggtggcag 1680
taggaccacc agtgtgggggt tctgctggga ccttgagagag cctgcatccc aggatgctgg 1740
gtggccctgc agcctcctcc acctcacctc catgacagcg ctaaacgttg gtgacggttg 1800
ggagcctctg gggctgttga agtcaccttg tgtgttccaa gtttccaaac aacagaaagt 1860
cattccttct ttttaaaatg gtgcttaagt tccagcagat gccacataag gggtttgcca 1920
tttgataccc ctggggaaca tttctgtaaa taccattgac acatccgcct tttgtataca 1980
tcctgggtaa tgagaggtgg cttttgcggc cagtattaga ctggaagttc atacctaatg 2040
actgtaataa tacctcaatg tttgaggagc atgttttgta taaaaatata ttgttaatct 2100
ctgttatgta ctgtactaat tcttacactg cctgtatact ttagtatgac gctgatacat 2160
aactaaattt gatacttata ttttcgtatg aaaatgagtt gtgaaagttt tgagtagata 2220
ttactttatc actttttgaa ctaagaaact tttgtaaaga aatttactat atatatatgc 2280
ctttttccta gcctgtttct tcctgttaat gtatttggtc atgtttggtg catagaactg 2340
ggtaaagca agtttctgtg ttttaatttct tcaaaatgta tatatttagt gctgcatctt 2400
atagcacttt gaaatacctc atgtttatga aaataaatag cttaaaatta aaaaaaaaaa 2460
aaaaaaa 2467

```

&lt;210&gt; 12

&lt;211&gt; 762

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 12

```

taccattctt caagaaacgg tttgaatcag actgcctttc cttttgtctt cattgtcata 60
aacatctgcc cccgtgtggg tctgactggc cgcaacccc taccgaagc ttttattcca 120
tcattgtgca ccgttgggtg ggaatgctgt ggcaacaggc cacgcctcca cttactggtt 180
ggctttgcgc aggcgccaac ggaagtgggt cgcaggaaga ggaagtcccg cctctctctc 240
ctcaggcagc agcaacgcgg aggaacaggg agtgaacgga gagcgtagtg accatcatga 300
gcctcctcaa caagcccaag agtgagatga cccagagga gctgcagaag cgagaggagg 360
aggaatttaa caccggtcca ctctctgtgc tcacacagtc agtcaagaac aatacccaag 420
tgctcatcaa ctgccgaac aataagaaac tcctgggccg cgtgaaggcc ttcgataggc 480
actgcaacat ggtgctggag aacgtgaagg agatgtggac tgaggtagcc aagagtggca 540
agggcaagaa gaagtccaag ccagtcaaca aagaccgcta catctccaag atgttcctgc 600
gcggggactc agtcatcgtg gtcctgcgga acccgctcat cgccggcaag taggggccgc 660
ctgtctgttg acagaactca ctctctgtc ctatgaagac cgctgccatt ggtgttgaga 720

```



ataataaagc tctgtgtttt tttctaaaaa aaaaaaaaaa aa

762

<210> 13

<211> 3379

<212> DNA

<213> Homo sapiens

<400> 13

aattccgcgg aatcatcgga atccttcacc atggcatcca gcccggccca gcgtcggcga	60
ggcaatgata ctctcacctc cagccctggc cgaagctccc ggcgtactga tgccctcacc	120
tccagccctg gccgtgacct tccaccattt gaggatgagt ccgaggggct cctaggcaca	180
gagggggccc tggaggaaga agaggatgga gaggagctca ttggagatgg catggaaagg	240
gactaccgcg ccatcccaga gctggacgcc tatgaggccg agggactggc tctggatgat	300
gaggacgtag aggagctgac ggccagtcga agggaggcag cagacgggcc atgcggcacg	360
gtgaccggga gctggccggg gctgggcgca tgcgccgtgg gctcctgtat gacagcgatg	420
aggaggacga ggagcgccct gcccgcaagc gccgccagtg gagccggcac ggaggacggc	480
gaggaggacg agcagatgat tgagagcatc gagaacctgg aggatctcaa aggccactct	540
gtgcgcgagt gggtagcat ggcgggcccc cggctggaga tccaccaccg cttcaagaac	600
ttcctgcgca ctacgctcga cagccacggc cacaacgtct tcaaggagcg catcagcgac	660
atgtgcaaag agaaccgtga gagcctggtg gtgaactatg aggacttggc agccagggag	720
cacgtgctgg cctacttcct gcctgaggca ccggcggagc tgctgcagat ctttgatgag	780
gctgccctgg aggtggtact ggccatgtac cccaagtacg accgcatcac caaccacatc	840
catgtccgca tctccacct gcctctggtg gaggagctgc gctcgtgag gcagctgcat	900
ctgaaccagc tgatccgcac cagtggggtg gtgaccagct gactggcgt cctgccccag	960
ctcagcatgg tcaagtacaa ctgcaacaag tgcaatttcg tcctgggtcc tttctgccag	1020
tcccagaacc aggaggtgaa accaggctcc tgtcctgagt gccagtcggc cggccccctt	1080
gaggtcaaca tggaggagac catctatcag aactaccagc gtatccgaat ccaggagagt	1140
ccaggcaaaag tggcggctcg gcggctgccc cgctccaagg acgccattct cctcgcagat	1200
ctggtggaca gctgcaacgc aggagacgag atagagctga ctggcatcta tcacaacaac	1260
tatgatggct ccctcaacac tgccaatggc ttccctgtct ttgccactgt catcctagcc	1320
aaccacgtgg ccaagaagga caacaagggt gctgtagggg aactgaccga tgaagatgtg	1380
aagatgatca ctagcctctc caaggatcag cagatcggag agaagatctt tgccagcatt	1440
gctccttcca tctatggtca tgaagacatc aagagaggcc ctgctctggc cctgttcgga	1500
ggggagccca aaaaccagg tggcaagcac aaggtacgtg gtgatatcaa cgtgctcttg	1560
tgccggagacc ctggcacagc gaagtcgcag tttctcaagt atattgagaa agtgtccagc	1620
cgagccatct tcaccactgg ccagggggcg tcggctgtgg ccgtcacggc gtatgtccag	1680

cggcaccctg tcagcagga gtggacctg gaggctggg ccctggttct ggctgaccga 1740  
 ggagtgtgtc tcattgatga atttgacaag atgaatgacc aggacagaac cagcatccat 1800  
 gaggccatgg agcaacagag catctccatc tcgaaggctg gcatcgtcac ctccctgcag 1860  
 gctcgtgca cggtcattgc tgccgccaac cccataggag ggcgctacga cccctcgtg 1920  
 actttctctg agaacgtgga cctcacagag cccatcatct cacgctttga catcctgtgt 1980  
 gtggtgaggg acaccgtgga cccagtccag gacgagatgc tggcccgtt cgtggtgggc 2040  
 agccacgtca gacaccaccc cagcaacaag gaggaggagg ggctggccaa tggcagcgt 2100  
 gctgagcccg ccatgccc aa cagtatggc gtggagcccc tgcccagga ggtcctgaag 2160  
 aagtacatca tctacgcaa ggagagggc caccgaagc tcaaccagat ggaccaggac 2220  
 aagggtggca agatgtacag tgacctgagg aaagaatcta tggcgacagg cagcatcccc 2280  
 attacggtgc ggcacatcga gtccatgagt catggcggag gccacgcgc gcatccatct 2340  
 gcgggactat gtgatcgaag acgacgtcaa catggccatc cgcgtgatgc tggagagctt 2400  
 catagacaca cagaagtcca gcgtcatcgc agcatgcga agacttttgc ccgctacctt 2460  
 tcattccggc gtgacaacaa tgagctgttg ctcttcatac tgaagcagtt agtggcagag 2520  
 caggtgacat atcagcga cgcgtttggg gccagcagg acactattga ggtccctgag 2580  
 aaggacttgg tgataaggc tcgtcagatc aacatccaca acctctctgc attttatgac 2640  
 agtgagctct tcaggatgaa caagttcagc cagcactga aaaggaaaat gatcctgag 2700  
 cagttctgag gccctatgcc atccataagg attccttggg attctggttt ggggtggtca 2760  
 gtgccctctg tgctttatgg acacaaaacc agagcacttg atgaactcgg ggtactaggg 2820  
 tcagggttta tagcaggatg tctggctgca cctggcatga ctgtttgttt ctccaagcct 2880  
 gctttgtgct tctcaccttt ggggtggatg ccttgccagt gtgtcttact tggttgctga 2940  
 acatcttgcc acctccgagt gctttgtctc cactcagtac cttggatcag agctgctgag 3000  
 ttcaggatgc ctgcgtgtgg tttaggtgtt agccttctta catggatgtc aggagagctg 3060  
 ctgccctctt ggcgtgagtt gcgtattcag gctgcttttg ctcgctttgg ccagagagct 3120  
 ggttgaagat gtttgtaatc gttttcagtc tcctgcagggt ttctgtgccc ctgtggtgga 3180  
 agaggcacga cagtgccagc gcagcgttct gggctcctca gtcgcagggg tgggatgtga 3240  
 gtcatgcgga ttatccactc gccacagtta tcagctgcca ttgctccctg tctgtttccc 3300  
 cactctctta tttgtgcatt cggtttggtt tctgtagttt taatttttaa taaagttgaa 3360  
 taaaatataa aaaaaaaaaa 3379

<210> 14

<211> 1488

<212> DNA

<213> Homo sapiens

<400> 14  
gttgggtgagc atcatggcaa ccgttacagc cacaaccaa gtcccggaga tccgtgatgt 60  
aacaaggatt gagcgaatcg gtgcccactc ccacatccgg ggactggggc tggacgatgc 120  
cttggagcct cggcaggctt cgcaaggcat ggtgggtcag ctggcggcac ggcgggcggc 180  
tggcgtggtg ctggagatga tccgggaagg gaagattgcc ggtcgggcag tccttattgc 240  
tggccagccg ggcacgggga agacggccat cgccatgggc atggcgcagg ccctgggccc 300  
tgacacgcca ttcacagcca tcgcccggcag tgaaatcttc tccctggaga tgagcaagac 360  
cgaggcgctg acgcaggcct tccggcggtc catcggcggt cgcatcaagg aggagacgga 420  
gatcatcgaa ggggaggtgg tggagatcca gattgatcga ccagcaacag ggacgggctc 480  
caaggtgggc aaactgacct tcaagaccac agagatggag accatctacg acctgggcac 540  
caagatgatt gagtccctga ccaaggacaa ggtccaggcc ggggacgtga tcaccatcga 600  
caaggcgacg ggcaagatct ccaagctggg ccgctccttc acacgcgccc gcgactacga 660  
cgctatgggc tcccagacca agttcgtgca gtgcccagat ggggagctcc agaaacgcaa 720  
ggaggtggtg cacaccgtgt ccctgcacga gatcgacgtc atcaactctc gcacccaggg 780  
cttcctggcg ctcttctcag gtgacacagg ggagatcaag tcagaagtcc gtgagcagat 840  
caatgccaa ggtggctgagt ggcgcgagga gggcaaggcg gagatcatcc ctggagtgtc 900  
gttcatcgac gaggtccaca tgctggacat cgagagcttc tccttcctca accgggcccct 960  
ggagagtgc atggcgcctg tcctgatcat ggccaccaac cgtggcatca cgcgaatccg 1020  
gggcaccagc taccagagcc ctcacggcat ccccatagac ctgctggacc ggctgcttat 1080  
cgtctccacc accccctaca gcgagaaaga cacgaagcag atcctccgca tccggtgcga 1140  
ggaagaagat gtggagatga gtgaggacgc ctacacggtg ctgacccgca tcgggctgga 1200  
gacgtcactg cgctacgcca tccagctcat cacagctgcc agcttgggtg gccggaaacg 1260  
caaggttaca gaagtgcagg tggtatgacat caagcgggtc tactcactct tcctggacga 1320  
gtcccgtcc acgcagtaca tgaaggagta ccaggacgcc ttctcttca acgaactcaa 1380  
aggcgagacc atggacacct cctgagttgg atgtcatccc ccgacccac cctgttttcc 1440  
accagagttc tgacactgtg actctgtata aaatggttg gaagctgc 1488

<210> 15

<211> 1811

<212> DNA

<213> Homo sapiens

<400> 15  
ggtttgtgta gagaggcgtg cagagcccgt tgtccggagt gcacctgtg cctgttctgt 60  
ccctcccggg agccccgcgc gctgtcgccc tcgagtcgcc atggaagtgc agaaagaggc 120  
acagcgcac atgaccctgt cgggtgtggaa gatgtatcac tcccgcacgc agcgcggtgg 180  
cctgcggctg caccggagtc tgacagctgtc gctggtcatg cgcagcggcc gggagctcta 240

```

cctctcggcc aaggtggagg ccctcgagcc cgaggtgtcg ttgccggccg ccctcccctc 300
tgaccctcgc ctgcaccgcg cccgagaagc cgagtccacg gccgagacag cgacccccga 360
cgggtagcac ccgtttccgg agccaatgga cacgcaggag gcgccgacag ccgaggagac 420
ctccgcctgc tgtgccccgc gccccgcaa agtcagccgc aaacgacgca gcagcagcct 480
gagcgacggc ggggacgttg gactggtccc gagcaagaaa gcccgctctg aagaaaagga 540
agaagaggag ggagcgtcat ccgaagtcgc cgatcgccctg cagccccctc cgggccaagc 600
ggagggcgcc tttcccaacc tggcccgcgt cctgcagagg cgcttctccg gcctcctgaa 660
ctgcagcccc gcggccccctc cgacggcgcc gccgcgtgc gaggcaaagc ccgcttgccg 720
cccgccggac agcatgctca acgtgctcgt gcgggccgtg gtggccttct gaggaccccg 780
agcggcgctg ccggagccca gagcgcgcgt cgaaccgtcg gcccgagggc gcagacctga 840
ggcgaggcca cccccctcca tcctggggga agcgcgccg aaaaccgtgg agagaagccg 900
ccgcccgggc tgctgagagg cccggagagg actctgtccc cggggagcca tcgccttcag 960
tgtgcaggga cggcaccgag gagtctgagc cgggcgcggg cgccttccgc agagacctgc 1020
gcccacaggt gctgtcttag tggactggga cgtgaacctt tcgctctcct tctggactgg 1080
gagaagggag gcttgggtgt tgtgtttttt gttttgtttg tttgtttgtt tttaaagatc 1140
tcctcagggt cggacttcat tttgtactgt gggctgtgct ggccctttca aggtttttca 1200
agagttggtt ttgcgtttcc aacctcgag aattccaggc actccccttc cccctccgct 1260
gacatacttg tataagcggg catcggtgcg tcatggggca ggcgtgggga gcttcctgtc 1320
gccttggctg ggtgtgggcc tggaggaagg tcctggggcg tgcaactcgc tgggcagtgg 1380
ggaggagagt ggcctgagtt acttcacccc cgcgtgctgc tggttaatgt cccgcgtctc 1440
tgcaccttcg ggtgggagcg gggactgatc tactttcaca ttctcaagtt tttctcatct 1500
gcattagagg tccccagtag gtccccaggt tccagcgtgc ccctccctca gacacacgga 1560
cacaatcagc cgagaagttc ctggtctgaa tcacgagaat gtggaggggt ggggggtgtc 1620
agtggaaagg cataaggctg agctgagacc agttgctggt gaaactgggc caatctgggg 1680
aggggaacat ccttgccagg gagtttctga gggctgctt tgtttacctt tcgtgcggtg 1740
gattcttttt aactccgtct acctggcggt ttgttagaaa tgtcagatag gaaaataaaa 1800
accatttgag t 1811

```

&lt;210&gt; 16

&lt;211&gt; 2038

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 16

```

ggcccgcggg actcagacca gcggggagcg cggcctccgc ccttggggcc ctcccgcggg 60
gccggagacc caagccccc acgccaggcc ctgccctgga agcgtcgcg gcccgccgcc 120

```

tggacggggg	agttgctgct	ctttggcgta	aattgcaatc	gattagggat	cgtttctcag	180
aatcaagtta	gaagtgagag	ttcagataag	tgaggccgcc	attgctgctt	tgaacacctc	240
agaaggggag	aatggattta	tcaggagtga	aaaagaagag	cttgctagga	gtcaaagaaa	300
ataataaaaa	gtccagcact	agggctcctt	cacctaccaa	acgcaaagac	cgctcagatg	360
agaagtccaa	ggatcgctca	aaagataaag	gggccaccaa	ggagtcgagt	gagaaggatc	420
gcggccggga	caaaacccga	aagaggcgca	gcgcttcag	tggtagcagc	agtaccaggt	480
ctcgggtccag	ctcgacttcc	agctcaggct	ccagcaccag	cactggctca	agcagtggct	540
ccagctcttc	ctcagcatcc	agccgctcag	gaagctccag	cacctcccgc	agctccagct	600
ctagcagctc	ttctggctct	ccaagtcctt	ctcggcgag	acacgacaac	aggaggcgct	660
cccgtccaa	atccaaacca	cctaaaagag	atgaaaagga	gaggaaaagg	cggagcccat	720
ctcctaagcc	caccaaagt	cacattggga	gactcaccg	gaatgtgaca	aaggatcaca	780
tcattggagat	attttccacc	tatgggaaaa	ttaaaatgat	tgacatgcc	gtggaaagga	840
tgcatcccca	tctgtccaaa	ggctatgcgt	acgtagagtt	tgagaatcca	gatgaagccg	900
agaaggcgct	gaagcacatg	gatggaggac	aaattgatgg	ccaggagatc	actgccaccg	960
ccgtgctggc	cccctggcct	aggccacccc	ccaggagatt	cagccctccc	aggagaatgt	1020
tgccaccacc	gcctatgtgg	cgaggtctc	ccccacggat	gaggagaagg	tcccgtccc	1080
cgaggcgag	gtccccgtg	cgccggagat	cacgggtccc	gggcccgcgc	cgccacagga	1140
gccgtccag	ctccaactcc	tcccgataaa	caggccactg	aagctctcgc	ccctgtaact	1200
tataccccac	ccagctcagt	tttgtcactt	ttctagccaa	aggaagacca	gtaggaaagc	1260
aaacccttga	ctctggcagg	atttgcaggc	agcaggcagc	acccctctgc	cagccggggc	1320
ccggctgca	aagtgtgtt	ggtttggatg	ctgtgtgcct	gtcaagattc	cctccggttt	1380
tctggctaga	aagctcatcc	gtttccggtt	tctaagagtc	agttcagtgg	cagagccacc	1440
agggaaaagt	gaggctcttg	gggttggttt	gaccctgctt	acctgggagc	acacttttcc	1500
cttccccgat	gacctgggat	ggtggccagg	ccgtgccctt	gctgttgctg	ggcagtgtcc	1560
ttttggaaag	ggagctgccc	caggcttttag	tgagctgcc	aaccctgtta	ggcctggcct	1620
ctcgaggcct	cttctgacct	caagggtcac	acccctcaa	agatcctctc	acccatggta	1680
gttgctgtc	gtggttctgt	ctgtccgtgc	accgatgcac	acaccgcacc	ccaccactgt	1740
actctgaaat	tggcgagtga	gtggagagcc	agctctgcgg	agtcatcacg	cagccatggt	1800
tgtgcctgcc	gttcatggtg	gtctttcagg	ttatcttggc	aacatgtaca	ttgcttttat	1860
tttttttctt	ttttgctttc	attgtacagt	cagtactata	aaatttctct	tttgagtttt	1920
atacctttgt	agcatttttag	atgacattgt	gtttgtactt	tgttgtgtag	agtggaagaa	1980
ttgtgttgaa	taaacccaag	atcggaatgc	aaaaaaaaa	aaaaaaaaa	aaaaaaaaa	2038

&lt;210&gt; 17

&lt;211&gt; 2062

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 17

```

gtcagtcctt cctgtagccg ccgccgccgc cggccgccgc ccctctgccg gcagctccgg      60
cgccacctcg ggcggcgctc tccggcgggc gggagccagg cgctgacggg cgcggcgggg      120
gcggccgagc gtcctgcgg ctgcgactca ggctccggcg tctgcgcttc cccatggggc      180
tggcctgcgg cgcctgggcg ctctgagatt gtcactgctg ttccaagggc acacgcagag      240
ggatttgaa ttcctggaga gttgcctttg tgagaagctg gaaatatttc tttcaattcc      300
atctcttagt tttccatagg aacatcaaga aatcatgaac aactttggta atgaagagtt      360
tgactgccac ttcctcgatg aaggttttac tgccaaggac attctggacc agaaaattaa      420
tgaagtttct tcttctgatg ataaggatgc cttctatgtg gcagacctgg gagacattct      480
aaagaaacat ctgaggtggt taaaagctct ccctcgtgtc accccctttt atgcagtcaa      540
atgtaatgat agcaaagcca tcgtgaagac cctgtctgct accgggacag gatttgactg      600
tgctagcaag actgaaatac agttggtgca gagtctgggg gtgcctccag agaggattat      660
ctatgcaaat ccttgtaaac aagtatctca aattaagtat gctgctaata atggagtcca      720
gatgatgact tttgatagtg aagttgagtt gatgaaagtt gccagagcac atcccaaagc      780
aaagttggtt ttgcggattg ccactgatga ttccaaagca gtctgtcgtc tcagtgtgaa      840
attcgggtgcc acgctcagaa ccagcaggct ctttttgaa cgggcgaaag agctaaatat      900
cgatgttggt ggtgtcagct tccatgtagg aagcggctgt accgatcctg agaccttcgt      960
gcaggcaatc tctgatgcc gctgtgtttt tgacatgggg gctgaggttg gtttcagcat      1020
gtatctgctt gatattggcg gtggctttcc tggatctgag gatgtgaaac ttaaatttga      1080
agagatcacc ggcgtaatca acccagcgtt ggacaaatac tttccgtcag actctggagt      1140
gagaatcata gctgagcccg gcagatacta tgttgcatca gctttcacgc ttgcagttaa      1200
tatcattgcc aagaaaattg tattaagga acagacgggc tctgatgacg aagatgagtc      1260
gagtgagcag acctttatgt attatgtgaa tgatggcgtc tatggatcat ttaattgcat      1320
actctatgac cagcacatg taaagcccct tctgcaaaag agacctaaac cagatgagaa      1380
gtattattca tccagcatat ggggaccaac atgtgatggc ctcgatcgga ttgttgagcg      1440
ctgtgacctg cctgaaatgc atgtgggtga ttggatgctc tttgaaaaca tgggcgctta      1500
cactgttgct gctgcctcta cgttcaatgg cttccagagg ccgacgatct actatgtgat      1560
gtcagggcct gcgtggcaac tcatgcagca attccagaac cccgacttcc caccgaagt      1620
agaggaacag gatgccagca ccctgcctgt gtcttggtcc tgggagagtg ggatgaaacg      1680
ccacagagca gcctgtgctt cggctagtat taatgtgtag atagcactct ggtagctgtt      1740
aactgcaagt ttagcttgaa ttaagggatt tggggggacc atgtaactta attactgcta      1800
gttttgaaat gtctttgtaa gagtaggggt gccatgatgc agccatatgg aagactagga      1860
tatgggtcac acttatctgt gttcctatgg aaactatttg aatatttggt ttatatggat      1920

```

ttttattcac tcttcagaca cgctactcaa gagtgtccct cagctgctga acaagcattt	1980
gtagcttgta caatggcaga atggggccaaa agcttagtgt tgtgacctgt ttttaaata	2040
aagtatcttg aaataattag gc	2062

&lt;210&gt; 18

&lt;211&gt; 2989

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<400> 18	
aattcgggca cgagggctct ccctccgcag cagccgagcc ggacctgcct ccccgggcgt	60
gctccgccgg ccccgccgcc ggcccgagc gacagacagg cgctccccgc agctccgcac	120
gggacctcagg ccgcccggacc ccagcgccgg accaccctct gtccgccccg aggagtttgc	180
cgctgcccgg agcacctgcg cacagatgga gctggaccac cggaccagcg gcgggctcca	240
cgctacccc gggccgcggg gcgggcaggt ggccaagccc aacgtgatcc tgcagatcgg	300
gaagtgccgg gccgagatgc tggagcacgt gcggcgagc caccggcacc tgcaggccga	360
ggtgtccaag caggtggagc gcgagctgaa ggggctgcac cggtcggtcg ggaagctgga	420
gagcaacctg gacggctacg tgcccacgag cgactcgag cgctggaaga agtccatcaa	480
ggcctgcctg tgccgctgcc aggagaccat cgccaacctg gagcgctggg tcaagcgcgga	540
gatgcacgtg tggcgcgagg tgttctaccg cctggagcgc tgggcccacc gcctggagtc	600
cacggggcggc aagtaccggg tgggcagcga gtcagccgc cacaccgttt ccgtgggcgt	660
gggggggtccc gagagctact gccacgaggc agacggctac gactacaccg tcagccccta	720
cgccatcacc ccgccccag ccgctggcga gctgccggg caggagccc cggaggcca	780
gcagtaccag ccgtgggtcc ccggcgagga cgggcagccc agccccggcg tggacacgca	840
gatcttcgag gaccctcgag agttcctgag ccacctagag gactacttgc ggcaggtggg	900
cggctctgag gactactggc tgtcccagat ccagaatcac atgaacgggc cggccaagaa	960
gtggtgggag ttcaagcagg gctccgtgaa gaactgggtg gagttcaaga aggagttcct	1020
gcagtacagc gagggcacgc tgtcccgaga ggccatccag cgggagctgg acctgccgca	1080
gaagcagggc gagccgctgg accagttcct gtggcgcaag cgggacctgt accagacgct	1140
ctacgtggac gcggacgagg aggagatcat ccagtacgtg gtgggcaccc tgcagcccaa	1200
gctcaagcgt ttctgcgcc accccctgcc caagacctg gagcagctca tccagagggg	1260
catggaggtg caggatgacc tggagcaggc ggccgagccg gccggcccc acctcccgt	1320
ggaggatgag gcggagaccc tcacgcccgc cccaacagc gagtccgtgg ccagtgaccg	1380
gaccagccc gagtagaggg catcccggag ccccgacct gccactaca tccagcctgt	1440
ggctttgccc accaggactt ttgagctggg gctgactcct gcaggggaag ccctggtcca	1500
gctgggtgcc ccctcgagct ccgggcggac tcgcacacac tcgtgtcatc cagatgtgag	1560

caccgcaccc agcggcaaag agccctcccc cctgcagggc tccacccatc accctccctc 1620  
 cgtctgtctt tccggcctgg accccaccct ccacactctc aggccatcac agaacacccc 1680  
 agcttcctca ttctgctaca acaccaggc cctctggaca tccagaaaac caagtgtccg 1740  
 gatggcaggg gccagcggcc accaagctca tgggacaccc agagcagaag ctagggcaga 1800  
 gccaatgctg agggagcctc gacttccggc gccgccgccc tctcccggca tccgcagagc 1860  
 cagctgacgc cctccctgcc tcccagggca gctggccagc ctccggcagc gcggccccct 1920  
 cctcccaggg gagagtagaa gtcgcacacg cagcagagca gacctgatgt cccggtgctt 1980  
 cctggccccct cagctccagt gattcacgcc cgcctggaga agaatacagag ctccagctcat 2040  
 gactcaccca tggcagggcg aggggtcccag aggggctgag tcctcaaata cggtgaggc 2100  
 agcagctggc accatcagag ccaggagagt gacaacaggc ctcaagggtc ccacaaagtc 2160  
 ttgtgctgctg tgctgggcac caccaccccc tcacctgca ggctgcctgc gtgggaggcg 2220  
 aagtcccagg acagcccaga ggggggctac agagaggagt cggtgcagc agagggcagg 2280  
 agcccagct tagccctgag cgccagcgcg aggaccaggg cctgccacta agcccgcccc 2340  
 gctggccgcc agctgcccgt cccagagcc actgcagcag gagtccggcc ctgcctccct 2400  
 cccagcaggg aaaccccgcc cgctgccagg ccatcctctc tgccagaggc tttcatgagc 2460  
 cccaaggctg gggccacagc tcctaccctt gccagcagc cctgagctca gctgcaggaa 2520  
 ggacatccca gaagccatgg ctctggggc gcttcaggc attctgccct gccccgacac 2580  
 cagaaccctg gtgctggtgg gccactagc tctgcagcct aagcagggtc tggctcaggg 2640  
 ttcacgttc tgcttctcc actgggggac cagccctgca gaccactctg acaagtcttc 2700  
 agcccacacc ctgccagccc cacagatttt atttttgcac ataagccata accaatcctc 2760  
 aaggctggca caggcttttg ggaagccctg gagcctgtga agaccctgga aacctcatga 2820  
 ggctgtggcc aaccctgccc ccttggccca cacagaccag gccttaaatg tcggtccagg 2880  
 ccctgtgcac cttaccccag agacagactc tttttgtaag atttttgttaa taaaacactg 2940  
 aaacttcaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2989

<210> 19

<211> 2365

<212> DNA

<213> Homo sapiens

<400> 19

gaaacggccc gagaagctcg cccggagaac ggggaggaat atgctgtgga gctcctctgc 60  
 catataaaca aaaagaggaa atctttcaaa catggctgaa gcaaagaccc actggcttgg 120  
 agcagccctg tctcttatcc ctttaatttt cctcatctct ggggctgaag cagcttcatt 180  
 tcagagaaac cagctgcttc agaaagaacc agacctcagg ttggaaaatg tccaaaagtt 240  
 tcccagtcct gaaatgatca gggcttttga gtacatagaa aacctccgac aacaagctca 300



taaggaagaa agcagcccag attataatcc ctaccaaggt gtctctgtcc cccttcagca	360
aaaagaaaat ggcgatgaaa gccacttgcc cgagagggat tctactgagtg aagaagactg	420
gatgagaata atactcgaag ctttgagaca ggctgaaaat gagcctcagt ctgcaccaaa	480
agaaaataag ccctatgcct tgaattcaga aaagaacttt ccaatggaca tgagtgatga	540
ttatgagaca cagcagtggc cagaaagaaa gcttaagcac atgcaattcc ctccctatgta	600
tgaagagaat tccagggata acccctttaa acgcacaaat gaaatagtgagggaacaata	660
tactctcaa agccttgcta catttgaatc tgtcttccaa gagctgggga aactgacagg	720
accaaacaac cagaaacgtg agaggatgga tgaggagcaa aaactttata cggatgatga	780
agatgatatc tacaaggcta ataacattgc ctatgaagat gtggctcgggg gagaagactg	840
gaaccagta gaggagaaaa tagagagtca aaccaggaa gaggtgagag acagcaaaga	900
gaatatagga aaaaatgaac aaatcaacga tgagatgaaa cgctcagggc agcttggcat	960
ccaggaagaa gatcttcgga aagagagtaa agaccaactc tcagatgatg tctccaaagt	1020
aattgcctat ttgaaaaggt tagtaaatgc tgcaggaagt gggagggttac agaattgggca	1080
aaatggggaa agggccacca ggctttttga gaaacctctt gattctcagt ctatttatca	1140
gctgattgaa atctcaagga atttacagat acccccagaa gacttaattg agatgctcaa	1200
aactggggag aagccgaatg gatcagtgga accggagcgg gagcttgacc ttcctgttga	1260
cctagatgac atctcagagg ctgacttaga ccatccagac ctgttccaaa ataggatgct	1320
ctccaagagt ggctacccta aaacacctgg tcgtgctggg actgaggccc taccagacgg	1380
gctcagtgtt gaggatattt taaatctttt agggatggag agtgcagcaa atcagaaaac	1440
gtcgtatttt cccaatccat ataaccagga gaaagtcttg ccaaggctcc cttatggtgc	1500
tggaagatct agatcgaacc agcttcccaa agctgcctgg attccacatg ttgaaaacag	1560
acagatggca tatgaaaacc tgaacgaca ggatcaagaa ttaggtgagt acttggccag	1620
gatgctagtt aaataccctg agatcattaa ttcaaaccaa gtgaagcgag ttcctggtca	1680
aggctcatct gaagatgacc tgcaggaaga ggaacaaatt gagcaggcca tcaaagagca	1740
tttgaatcaa ggcagctctc aggagactga caagctggcc ccggtgagca aaagggtccc	1800
tgtggggccc ccgaagaatg atgatacccc aaataggcag tactgggatg aagatctgtt	1860
aatgaaagtg ctggaatacc tcaatcaaga aaaggcagaa aagggaaggg agcatattgc	1920
taagagagca atggaaaata tgtaagctgc ttctattaat taccctactt tcattcctcc	1980
cacccaagc aaatcccaac atttctcttc agtgtgttga cttctatcct gttaacactg	2040
taatattctt aaatgatgta caggcagatg aaaccaggtc actggggagt ctgcttcatt	2100
tcctctgagc tgttatcttg tgtatggata tgtgtaaaatg ttatgactcc ttgataaaaa	2160
atttattatg tccattattc aagaaagata tctatgactg tgtttaatag tatatcta	2220
ggctgtggca ttgttgatgc tcacatatga taaaaaagtg tcctataatt ctattgaaag	2280
tttttaatat ttattgaatt attttggttac tgtctgtagc gttttgtgga gtactggacc	2340
aaaaaataa agcattataa atata	2365

&lt;210&gt; 20

&lt;211&gt; 2825

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

&lt;400&gt; 20

```

gtacggcttc cggtggcggg acgcggggcc gcgcacgcgg gaaaagcttc cccggtgtcc      60
ccccatcccc ctccccgcgc cccccccgcg tccccccagc gcgcccacct ctcgcgccgg      120
ggccctcgcg aggccgcagc ctgaggagat tccaacctg ctgagcatcc gcacaccac      180
tcaggagttg gggcccagct cccagtttac ttggtttccc ttgtgcagcc tggggctctg      240
cccaggccac cacaggcagg ggtcgacatg gcagagacac tggagttcaa cgacgtctat      300
caggaggtga aaggttccat gaatgatggt cgactgaggt tgagccgtca gggcatcatc      360
ttcaagaata gcaagacagg caaagtggac aacatccagg ctggggagtt aacagaaggt      420
atctggcgcc gtgttgctct gggccatgga cttaactgc ttacaaagaa tggccatgtc      480
tacaagtatg atggcttccg agaatcgag tttgagaaac tctctgattt cttcaaaact      540
cactatcgcc ttgagctaata ggagaaggac ctttgtgtga agggctggaa ctgggggaca      600
gtgaaatttg gtgggcagct gctttccttt gacattggtg accagccagt ctttgagata      660
cccctcagca atgtgtccca gtgcaccaca ggcaagaatg aggtgacact ggaattccac      720
caaaacgatg acgcagaggt gtctctcatg gaggtgcgct tctacgtccc acccaccag      780
gaggatggtg tggaccctgt tgaggccttt gcccagaatg tgttgtaaaa ggcggatgta      840
atccaggcca cgggagatgc catctgcac ttcggggagc tgcagtgtct gactcctcgt      900
ggtcgttatg acattcggat ctacccacc tttctgcacc tgcattggcaa gacctttgac      960
tacaagatcc cctacaccac agtactgcgt ctgtttttgt taccacacaa ggaccagcgc     1020
cagatgttct ttgtgatcag cctggatccc ccaatcaagc aaggccaaac tcgctaccac     1080
ttcctgatcc tcctcttctc caaggacgag gacatttcgt tgactctgaa catgaacgag     1140
gaagaagtgg agaagcgctt tgagggtcgg ctcaccaaga acatgtcagg atccctctat     1200
gagatggtca gccgggtcat gaaagcactg gtaaaccgca agatcacagt gccaggcaac     1260
ttccaagggc actcaggggc ccagtgcatt acctgttcct acaaggcaag ctcaggactg     1320
ctctacccgc tggagcgggg cttcatctac gtccacaagc cacctgtgca catccgcttc     1380
gatgagatct cttttgtcaa ctttgctcgt ggtaccacta ctactcgttc ctttgacttt     1440
gaaattgaga ccaagcaggg cactcagtat accttcagca gcattgagag ggaggagtac     1500
gggaaactgt ttgattttgt caacgcgaaa aagctcaaca tcaaaaaccg aggattgaaa     1560
gagggtcatg acccaagcta cgatgaatat gctgactctg atgaggacca gcatgatgcc     1620
tacttgagga ggatgaagga ggaaggcaag atccgggagg agaattgcaa tgacagcagc     1680
gatgactcag gagaagaaac cgatgagtca ttcaaccag gtgaagagga ggaagatgtg     1740

```

gcagaggagt ttgacagcaa cgctctgcc agctcctcca gtaatgaggg tgacagtgac 1800  
 cgggatgaga agaagcggaa acagctcaaa aaggccaaga tggccaagga ccgcaagagc 1860  
 cgcaagaagc ctgtggaggt gaagaagggc aaagaccca atgccccaa gaggcccatg 1920  
 tctgcataca tgctgtggct caatgccagc cgagagaaga tcaagtcaga ccatcctggc 1980  
 atcagcatca cggatctttc caagaaggca ggcgagatct ggaaggggaat gtccaaagag 2040  
 aagaaagagg agtgggatcg caaggctgag gatgccagga gggactatga aaaagccatg 2100  
 aaagaatatg aagggggccg aggcgagtct tctaagaggg acaagtcaaa gaagaagaag 2160  
 aaagtaaagg taaagatgga aaagaaatcc acgccctcta ggggctcatc atccaagtcg 2220  
 tcctcaaggc agctaagcga gagcttcaag agcaaagagt ttgtgtctag tgatgagagc 2280  
 tcttcgggag agaacaagag caaaaagaag aggaggagga gcgaggactc tgaagaagaa 2340  
 gaactagcca gtactcccc cagctcagag gactcagcgt caggatccga tgagtagaaa 2400  
 cggaggaagg ttctctttgc gcttgccctc tcacaccccc cgactcccca cccatatttt 2460  
 ggtaccagtt tctcctcatg aaatgcagtc cctggattct gtgccatctg aacatgctct 2520  
 cctgttggtg tgtatgtcac tagggcagtg gggagacgtc ttaactctgc tgcttcccaa 2580  
 ggatggctgt ttataatttg gggagagata ggggtggagg cagggcaatg caggatccaa 2640  
 atcctcatct tactttccc accttaagga tgtagctgct gcttgcctg ttcaagttgc 2700  
 tggagcaggg gtcagtgtgag gccaggcctg tagctcctac ctggggccta tttctacttt 2760  
 cattttgtat ttctggctg tgaaaatgat ttaataaagg gaactgactt tggaaccaa 2820  
 aaaaa 2825

<210> 21

<211> 10488

<212> DNA

<213> Homo sapiens

<400> 21  
 aagagttttc ctccgcagct ctgagtctcc acttttttgg tggagaaagg ctgcaaaaag 60  
 aaaaagagac gcagtgagtg ggaaaagtat gcctcctatt caaacctaatt tgaatcgagg 120  
 agcccaggga cacacgcctt cagggttgct caggggttca tatttggtgc ttagacaaat 180  
 tcaaaatgag gaaacatcgg cacttgccct tagtggccgt cttttgcctc tttctctcag 240  
 gctttctac aactcatgcc cagcagcagc aagcagatgt caaaaatggg gcggctgctg 300  
 atataatatt tctagtggat tcctcttgga ccattggaga ggaacatttc caacttgttc 360  
 gagagtttct atatgatgtt gtaaaatcct tagctgtggg agaaaatgat ttccattttg 420  
 ctctggtcca gttcaacgga aacccacata ccgagttcct gttaaatacg tatcgtacta 480  
 aacaagaagt cttttctcat atttccaaca tgtcttatat tgggggaacc aatcagactg 540  
 gaaaaggatt agaatacata atgcaaagcc acctcaccaa ggctgctgga agccgggccc 600

gtgacggagt ccctcaggtt atcgtagtgt taactgatgg aactcgaag gatggccttg 660  
ctctgccctc agcggaactt aagtctgctg atgttaacgt gtttgcaatt ggagttgagg 720  
atgcagatga aggagcgta aaagaaatag caagtgaacc gctcaatatg catatgttca 780  
acctagagaa ttttacctca cttcatgaca tagtaggaaa cttagtgtcc tgtgtgcatt 840  
catccgtgag tccagaaagg gctggggaca cggaaccct taaagacatc acagcacaag 900  
actctgctga cattatcttc cttattgatg gatcaaaca caccggaagt gtcaatttcg 960  
cagtcattct cgacttcctt gtaaactctc ttgagaaact cccaattgga actcagcaga 1020  
tccgagtggg ggtggtccag tttagcgatg agcccagaac catgttttcc ttggacacct 1080  
actccacca ggcccaggtt ctgggtgcag tgaagccct cgggtttgct ggtggggagt 1140  
tggccaatat cggcctcgcc cttgatttcg tgggtggaga ccacttcacc cgggcagggg 1200  
gcagccgctg ggaggaagg gttccccagg tgctggctct cataagtgcc gggccttcta 1260  
gtgacgagat tcgctacggg gtggtagcac tgaagcaggc tagcgtgttc tcattcggcc 1320  
ttggagccca ggccgcctcc agggcagagc ttcagcacat agctaccgat gacaacttgg 1380  
tgtttactgt cccggaattc cgtagctttg gggacctcca ggagaaatta ctgccgtaca 1440  
ttgttgccgt ggcccaaagg cacattgtct tgaaaccgcc aaccattgtc acacaagtca 1500  
ttgaagtcaa caagagagac atagtcttcc tgggtggatg ctcatctgca ctgggactgg 1560  
ccaacttcaa tgccatccga gacttcattg ctaaagtcac ccagaggctg gaaatcggac 1620  
aggatcttat ccagggtggca gtggcccagt atgcagacac tgtgaggcct gaattttatt 1680  
tcaataccca tccaacaaaa agggaagtca taaccgctgt gcggaaaatg aagcccctgg 1740  
acggctcggc cctgtacacg ggctctgctc tagactttgt tcgtaacaac ctattcacga 1800  
gttcagccgg ctaccgggct gccgagggga ttcctaagct tttggtgctg atcacagggtg 1860  
gtaagtccct agatgaaatc agccagcctg ccaggagct gaagagaagc agcataatgg 1920  
cctttgccat tgggaacaag ggtgccgatc aggctgagct ggaagagatc gctttcgact 1980  
cctccctggt gttcatccca gctgagttcc gagccgcccc attgcaaggc atgctgcctg 2040  
gcttgctggc acctctcagg accctctctg gaaccctga agttcactca aacaaaagag 2100  
atatcatctt tcttttggat ggatcagcca acgttgga aaccaatttc cttatgtgc 2160  
gcgactttgt aatgaaccta gttaacagcc ttgatattgg aaatgacaat attcgtgttg 2220  
gttttagtgca atttagtgac actcctgtaa cggagtctc tttaaacaca taccagacca 2280  
agtcagatat ccttggtcat ctgaggcagc tgcagctcca gggagggttcg ggcctgaaca 2340  
caggctcagc cctaagctat gtctatgcc accacttcac ggaagctggc ggcagcagga 2400  
tccgtgaaca cgtgccgag ctctctcttc tgctcacagc tgggcagtct gaggactcct 2460  
atgtgcaagc tgccaacgcc ttgacacgcg cgggcacct gactttttgt gtgggagcta 2520  
gccaggcgaa taaggcagag cttgagcaga ttgcttttaa cccaagcctg gtgtatctca 2580  
tggatgattt cagctccctg ccagctttgc ctgagcagct gattcagccc ctaaccacat 2640  
atgttagtgg aggtgtggag gaagtaccac tcgctcagcc agagagcaag cgagacattc 2700

tggttcctctt tgacggctca gccaatcttg tgggccagtt ccctgttggtc cgtgactttc 2760  
tctacaagat tatcgatgag ctcaatgtga agccagaggg gacccgaatt gcggtggctc 2820  
agtacagcga tgatgtcaag gtggagtccc gttttgatga gcaccagagt aagcctgaga 2880  
tcctgaatct tgtgaagaga atgaagatca agacgggcaa agccctcaac ctggggtacg 2940  
cgctggacta tgcacagagg tacatTTTTg tgaagtctgc tggcagccgg atcgaggatg 3000  
gagtgcctca gttcctgggtg ctgctgggtc caggaagggtc atctgaccgt gtggatgggc 3060  
cagcaagtaa cctgaagcag agtgggggtg tgcctttcat cttccaagcc aagaacgcag 3120  
accctgctga gttagagcag atcgtgctgt ctccagcgtt tatcctggct gcagagtcgc 3180  
ttcccaagat tggagatctt catccacaga tagtgaatct cttaaaatca gtgcacaacg 3240  
gagcaccagc accagtttca ggtgaaaagg acgtgggtgt tctgcttgat ggctctgagg 3300  
gcgtcaggag cggcttccct ctgttgaaag agtttgcca gagagtgggtg gaaagcctgg 3360  
atgtgggcca ggaccgggtc cgctggccg tgggtgcagta cagcgaccgg accaggcccg 3420  
agttctacct gaattcatat atgaacaagc aggacgtcgt caacgctgtc cgccagctga 3480  
ccctgctggg agggccgacc cccaacaccg gggccgccct ggagtttgtc ctgaggaaca 3540  
tcctggctcag ctctgcggga agcaggataa cagaagggtgt gcccagctg ctgatcgtcc 3600  
tcacggccga caggctctggg gatgatgtgc ggaacccctc cgtggctcgtg aagaggggtg 3660  
gggctgtgcc cattggcatt ggcctcggga acgctgacat cacagagatg cagaccatct 3720  
ccttcatccc ggactttgct gtggccattc ccaccttcg ccagctgggg accgtccaac 3780  
aggctcatctc tgagaggggtg acccagctca cccgcgagga gctgagcagg ctgcagccgg 3840  
tggtgcagcc tctaccgagc ccaggtgttg gtggcaagag ggacgtgggtc tttctcatcg 3900  
atgggtccca aagtgccggg cctgagttcc agtacgttcg caccctcata gagaggctgg 3960  
ttgactacct ggacgtgggc tttgacacca cccgggtggc tgtcatccag ttcagcgatg 4020  
acccaaggc ggagtctctg ctgaacgccc attccagcaa ggatgaagtg cagaacgcgg 4080  
tgcagcggct gagggccaag ggagggcggc agatcaacgt gggcaatgcc ctggagtacg 4140  
tgtccaggaa catcttcaag agggccctgg ggagccgcat tgaagagggc gtcccacagt 4200  
tcctggctct catctcgtct ggaaagtctg acgatgaggt ggtcgtcccg gcggtggagc 4260  
tcaagcagtt tggcgtggcc cctttcacga tcgccaggaa cgagaccag gaggagctgg 4320  
tgaagatctc gctgagcccc gaatatgtgt tctcgggtgag caccttccgg gagctgcca 4380  
gcctggagca gaaactgctg acgcccata cagacctgac ctgagagcag atccagaagc 4440  
tcttagccag cactcgtat ccacctccag cagttgagag tgatgctgca gacattgtct 4500  
ttctgatcga cagctctgag ggagttaggc cagatggctt tgcaatatt cgagattttg 4560  
ttagcaggat tgttcgaaga ctcaacatcg gcccagtaa agtgagagtt ggggtcgtgc 4620  
agttcagcaa tgatgtcttc ccagaattct atctgaaaac ctacagatcc caggccccgg 4680  
tgctggacgc catacggcgc ctgaggctca ggggggggtc cccactgaac actggcaagg 4740  
ctctcgaatt tgtggcaaga aacctcttg ttaagtctgc ggggagtcgc atagaagacg 4800

gggtgcccc acacctgggt ctggctctgg gtggaaaatc ccaggacgat gtgtccaggt 4860  
tcgcccaggt gatccgttcc tcgggcattg tgagttagg ggtaggagac cggaacatcg 4920  
acagaacaga gctgcagacc atcaccaatg accccagact ggtcttcaca gtgcgagagt 4980  
tcagagagct tcccaacata gaagaaagaa tcatgaactc gtttggaacc tccgcagcca 5040  
ctcctgcacc tccaggggtg gacacccctc ctcttcacg gccagagaag aagaaagcag 5100  
acattgtgtt cctgttggat gggtccatca acttcaggag ggacagtttc caggaagtgc 5160  
ttcgttttgt gtctgaaata gtggacacag tttatgaaga tggcgactcc atccaagtgg 5220  
ggcttgtcca gtacaactct gacccactg acgaattctt cctgaaggac ttctctacca 5280  
agaggcagat tattgacgcc atcaacaaag tgggtctaca agggggaaga cacgccaaca 5340  
ctaagggtgg ccttgagcac ctgcggttaa accactttgt gcctgaggca ggcagccgcc 5400  
tggaccagcg ggtccctcag attgccttg tgatcacggg aggaaagtcg gtggaagatg 5460  
cacaggatgt gagcctggcc ctacccaga ggggggtcaa agtgtttgct gttggagtga 5520  
ggaatatcga ctcgaggag gttggaaaga tagcgtccaa cagcgccaca gcgttccgcg 5580  
tgggcaacgt ccaggagctg tccgaactga gcgagcaagt tttggaaact ttgcatgatg 5640  
cgatgcatga aaccctttgc cctggtgtaa ctgatgctgc caaagcttgt aatctggatg 5700  
tgattctggg gtttgatggt tctagagacc agaatgtttt tgtggcccag aagggttcg 5760  
agtccaaggt ggacgccatc ttgaacagaa tcagccagat gcacaggggtc agctgcagcg 5820  
gtggccgctc gccaccgtg cgtgtgtcag tgggtggcaa cacgccctcg ggcccgttg 5880  
aggcctttga ctttgacgag taccagccag agatgctcga gaagtccgg aacatgcgca 5940  
gccagaccc ctacgtcctc acggaggaca ccctgaaggt ctacctgaac aagttcagac 6000  
agtcctcgcc ggacagcgtg aagggtgtca ttcatTTTtac tgatggagca gacggagatc 6060  
tggctgattt acacagagca tctgagaacc tccgcaaga aggagtccgt gccttgatcc 6120  
tgggtggcct tgaacgagt gtcaacttg agcggtaat gcatctggag tttgggcgag 6180  
ggtttatgta tgacaggccc ctgaggctta acttgctgga cttggattat gaactagcgg 6240  
agcagcttga caacattgcc gagaaagctt gctgtggggt tccctgcaag tgctctgggc 6300  
agaggggaga ccgcgggccc atcggcagca tcgggcaaaa gggatttcct ggagaagacg 6360  
gctaccgagg ctatcctggt gatgaggggtg gacccggtga gcgtggtccg cctggtgtga 6420  
acggcactca aggtttccag ggctgcccgg gccagagagg agtaaagggc tctcggggat 6480  
tcccaggaga gaagggcgaa gtaggagaaa ttggactgga tggcttgat ggtgaagatg 6540  
gagacaaagg attgcctggt tcttctggag agaaaggga tcctggaaga aggggtgata 6600  
aaggacctcg aggagagaaa ggagaaagag gagatgttgg gattcgaggg gaccgggta 6660  
accaggaca agacagccag gagagaggac ccaaaggaga aaccggtgac ctcggcccca 6720  
tgggtgtccc agggagagat ggagtacctg gaggacctgg agaaactggg aagaatggtg 6780  
gctttggccg aaggggacct cccggagcta agggcaacaa gggcggtcct ggccagccgg 6840  
gctttgaggg agagcagggg accagaggtg cacagggccc agctggtcct gctggtcctc 6900

cagggctgat aggagaacaa ggcatttctg gacctagggg aagcggagggt gcccgtggcg 6960  
ctcctggaga acgaggcaga accggtccac tgggaagaaa gggtagagccc ggagagccag 7020  
gacaaaaagg aggaatcggg aacccgggcc ctctgtggga gacgggagat gacgggagag 7080  
acggagttag cagtgaagga cgcagaggca aaaaaggaga aagaggattt cctggatacc 7140  
caggaccaa gggtaaccca ggtgaacctg ggctaaatgg aacaacagga cccaaaggca 7200  
tcagaggccg aaggggaaat tcgggacctc cagggatagt tggacagaag gggagacctg 7260  
gctaccagc accagctggt ccaaggggca acaggggcca ctccatcgat caatgtgccc 7320  
tcaccaaag catcaaagat aaatgccctt gctgttacgg gcccctggag tgccccgtct 7380  
tcccaacaga actagccttt gctttagaca cctctgaggg agtcaaccaa gacactttcg 7440  
gccggatgag agatgtggc ttgagtattg tgaatgtcct gaccattgct gagagcaact 7500  
gcccagcggg gggccgggtg gctgtgtgca cctacaacaa cgagggtgacc acggagatcc 7560  
ggtttctga ctccaagagg aagtcggtcc tcctggacaa gattaagaac cttcaggtgg 7620  
ctctgacatc caaacagcag agtctggaga ctgccatgct gtttgtggcc aggaacacat 7680  
ttaagcgtgt gaggaacgga ttcctaata ggaagtggc tgttttcttc agcaacacac 7740  
ccacaagagc atccccacag ctacagaggg ctgtgctcaa actctcagat gcggggatca 7800  
cccccttgtt ccttacaagg caggaagacc ggcagctcat caacgccttg cagatcaata 7860  
acacagcagt ggggcatgag cttgtcctgc ctgcaggagg agacctcaca gacttcctgg 7920  
agaatgtcct cacgtgtcat gtttgcttgg acatctgcaa catcgacca tcctgtggat 7980  
ttggcagttg gaggccttcc ttcagggaca ggagagcggc agggagtgat gtggacatcg 8040  
acatggcttt catcttagac agcgtgaga ccaccacct gttccagttc aatgagatga 8100  
agaagtacat agcgtacctg gtcagacaac tggacatgag ccagatccc aaggcctccc 8160  
agcacttcgc cagagtggca gttgtgcagc acgcgcctc ttagtccgtg gacaatgcca 8220  
gcatgccacc tgtgaagggt gaattctccc tgactgacta tggctccaag gagaagctgg 8280  
tggacttcct cagcagggga atgacacagt tgcagggaac cagggcctta ggcagtgcca 8340  
ttgaatacac catagagaat gtctttgaaa gtgccccaaa cccacgggac ctgaaaattg 8400  
tggctcctgat gctgacgggc gaggtgccgg agcagcagct ggaggaggcc cagagagtca 8460  
tcctgcaggc caaatgcaag ggctacttct tcgtggctct gggcattggc aggaaggtga 8520  
acatcaagga ggtatacacc ttcgccagt agccaaacga cgtcttcttc aaattagtgg 8580  
acaagtccac cgagctcaac gaggagcctt tgatgcgctt cgggaggctg ttgccgtcct 8640  
tcgtcagcag tgaaaatgct ttttacttgt cccagatat caggaaacag tgtgattggt 8700  
tccaagggga ccaaccaca aagaacctg tgaagtttg tcacaaacaa gtaaattgtc 8760  
cgaataacgt tacttcaagt cctacatcca acccagtga gacaacgaag ccggtgacta 8820  
cgacgaagcc ggtgaccacc acaacaaagc ctgtaaccac cacaacaaag cctgtgacta 8880  
ttataaatca gccatctgtg aagccagccg ctgcaaagcc gggccctgcg aaacctgtgg 8940  
ctgccaagcc tgtggccaca aagacggcca ctgttagacc ccagtggcg gtgaagccag 9000

```

caacagcagc gaagcctgta gcagcaaagc cagcagctgt aagaccccc gctgctgctg 9060
caaaaccagt ggcgaccaag cctgagggtcc ctaggccaca ggcagccaaa ccagctgcca 9120
ccaagccagc caccactaag cccgtgggta agatgctccg tgaagtccag gtgtttgaga 9180
taacagagaa cagcgccaaa ctccactggg agaggcctga gcccccggt cttattttt 9240
atgacctcac cgtcacctca gcccatgata agtccctggg tctgaagcag aacctcacgg 9300
tcacggaccg cgtcattgga ggcctgctcg ctgggcagac ataccatgtg gctgtggtct 9360
gctacctgag gtctcagggt agagccacct accacggaag tttcagtaca aagaaatctc 9420
agccccacc tccacagcca gcaaggctcag cttctagtcc aaccatcaat ctaatggtga 9480
gcacagaacc attggtcttc actgaaacag atatatgcaa gttgccgaaa gacgaaggaa 9540
cttgaggga tttcatatta aaatggtact atgatccaaa caccaaaagc tgtgcaagat 9600
tctgggtatg aggttggtg ggaaacgaaa acaattttg atcacagaaa gaatgtgaaa 9660
aggtttgctc tcctgtgctc gccaaacccg gagtcacag tgtgatggga acctaagcgt 9720
gggtggccaa catcatatac ctcttgaaga agaaggagtc agccatcgcc aacttgctc 9780
tgtagaagct ccgggtgtag attcccttgc actgtatcat ttcattgctt gatttacact 9840
cgaactcggg aggggaacatc ctgctgcatg acctatcagt atggtgctaa tgtgtctgtg 9900
gacctcgtc ctctgtctcc agcagttctc tcgaatactt tgaatgttgt gtaacagtta 9960
gccactgctg gtgtttatgt gaacattcct atcaatccaa attccctctg gagtttcatg 10020
ttatgcctgt tgcaggcaaa tgtaaagtct agaaaataat gcaaatgtca cggctactct 10080
atatactttt gcttggttca ttttttttcc cttttagtta agcatgactt tagatgggaa 10140
gcctgtgtat cgtggagaaa caagagacca actttttcat tccctgcccc caatttcca 10200
gactagattt caagctaatt ttctttttct gaagcctcta acaaatgata tagttcagaa 10260
ggaagcaaaa tcccttaatc tatgtgcacc gttgggacca atgccttaat taaagaattt 10320
aaaaaagttg taatagagaa tatttttggc attcctctca atgttgtgtg tttttttttt 10380
ttgtgtgctg gagggagggg atttaatttt aattttaaaa tgtttaggaa atttatacaa 10440
agaaactttt taataaagta tattgaaagt ttaaaaaaaa aaaaaaaa 10488

```

<210> 22

<211> 1044

<212> DNA

<213> Homo sapiens

<400> 22

```

gaattccctg aggaggcgaa tccggcgggg atcagagcca tcagaaccgc caccatgacg 60
gtgggcaaga gcagcaagat gctgcagcat attgattaca ggatgagggtg catcctgcag 120
gacggccgga tcttcattgg caccttcaag gcttttgaca agcacatgaa tttgatcctc 180
tgtgactgtg atgagttcag aaagatcaag ccaaagaact ccaaacaagc agaaagggaa 240

```



gagaagcgag tcctcggctt ggtgctgctg cgaggggaga atctggtctc aatgacagta 300  
 gagggacctc ctcccaaaga tactgggtatt gctcgagttc cacttgctgg agctgccggg 360  
 ggcccaggga tcggcagggc tgctggcaga ggaatcccag ctggggttcc catgcccag 420  
 gctcctgcag gacttgctgg gccagtcctg ggggttggcg ggccatccca acaggtgatg 480  
 accccacaag gaagaggtag tggtgcagcc gctgcagctg ctgccacagc cagtattgcc 540  
 ggggctccaa cccagtaccc acctggccgt gggggctctc cccacctat gggccgagga 600  
 gcacccctc caggcatgat gggccacct cctggtatga gacctctat gggccccca 660  
 atggggatcc cccctggaag agggactcca atgggcatgc cccctccggg aatgcggcct 720  
 cctccccctg ggatgcgagg ccttctttga ccctggcca cagagtatgg aagtagctcc 780  
 gcagaggcgt gggctcgatt cctcagggcc acgttaccac agacctgttt gtttcttatg 840  
 ctgttgttcg tggagtctca tgggattgtc tggtttcct tacaggggcc cctccccgg 900  
 gaatgcgccc accaaggccc tagactcatc ttggccctcc tcagctccct gcctgtttcc 960  
 cgtaaggctg tacatagtcc ttttatctcc ttgtggccta tgaaactggt ttataataaa 1020  
 ctcttaagag aacattataa ttgc 1044

<210> 23

<211> 1475

<212> DNA

<213> Homo sapiens

<400> 23

gtcgacgcgg ccgcgctccg ctcccgtag taacttggct ccgggggctc cgctcgctg 60  
 cccgcacgcc gcccgccacc caggaccgag ccgcccgcct ccgcccgtag caaaccttc 120  
 cgacggccct cgctgcgcaa gccgggacgc ctctcccccc tccgcccccg ccgcggaaag 180  
 ttaagtttga agagggggga agaggggaac atggacatga agaggaggat ccacctggag 240  
 ctgaggaacc ggaccccgcc agctgttcga gaacttgtct tggacaattg caaatcaaat 300  
 gatggaaaaa ttgagggtt aacagctgaa tttgtgaact tagagttcct cagtttaata 360  
 aatgtaggct tgatctcagt ttcaaatctc cccaagctgc cttaaattgaa aaagcttgaa 420  
 ctgagtgaaa atagaatctt tggagggtctg gacatgttag ctgaaaaact tccaaatctc 480  
 acacatctaa acttaagtgg aaataaactg aaagatatca gcaccttga acctttgaaa 540  
 aagttagaat gtctgaaaag cctggacctc tttactgtg aggttaccaa cctgaatgac 600  
 taccgagaga gtgtcttcaa gctcctgccc cagcttacct acttgatgg ctatgaccga 660  
 gaggaccagg aagcacctga ctgagatgcc gaggtggatg gtgtggatga agaggaggag 720  
 gacgaagaag gagaagatga ggaagacgag gacgatgagg atggtgaaga agaggagttt 780  
 gatgaagaag atgatgaaga tgaagatgta gaaggggatg aggacgacga tgaagtcagt 840  
 gaggaggaag aagaatttgg acttgatgaa gaagatgaag atgaggatga ggatgaagag 900

gaggaagaag gtgggaaagg tgaaaagagg aagagagaaa cagatgatga aggagaagat 960  
 gattaagacc ccagatgacc tgcagaaaca gaactgttca gtattgggtg gactgctcat 1020  
 ggattttgta gctgtttaaa aaaaaaaaaa aggtagctgt gatacaaacc ccaggacacc 1080  
 caccaccca aagagccaaa gaatagtcc tgtgacattc cgccttcctt ccatgtagtc 1140  
 cctcttggtg atctaccacc aagcttggtg acttcacccc aacaaaattg taagcgttgt 1200  
 taggtttttg tgtaagattc ttgctgtagc gtggatagct gtgattgggtg agtcaaccgt 1260  
 ctgtggctac cagttacact gagattgtaa cagcattttt actttctgta caacaaaaaa 1320  
 gctttgtaaa taaaaatctta acatttttggg tctgtttttt catgctttgc tttttaatta 1380  
 ttattattat tttttttaca ttaggacatt ttatgtgaca actgccaaaa aagtattttt 1440  
 aagaatttaa gcgaataaaa cagttactct ttggc 1475

<210> 24

<211> 2690

<212> DNA

<213> Homo sapiens

<400> 24

gctcttttct cgggacggga gaggccgtgt agcgtcgccg ttactccgag gagataccag 60  
 tcggtagagg agaagtcgag gttagagggg actgggaggc actttgctgt ctgcaatcga 120  
 agttgagggg gcaaaaatgc agagtaataa aacttttaac ttggagaagc aaaaccatac 180  
 tccaagaaag catcatcaac atcaccacca gcagcagcac caccagcagc aacagcagca 240  
 gccgccacca ccgccaatac ctgcaaatgg gcaacaggcc agcagccaaa atgaaggctt 300  
 gactattgac ctgaagaatt ttagaaaacc aggagagaag accttcaccc aacgaagccg 360  
 tctttttgtg ggaaatcttc ctcccgacat cactgaggaa gaaatgagga aactatttga 420  
 gaaatatgga aaggcaggcg aagtcttcat tcataaggat aaaggatttg gctttatccg 480  
 cttggaaacc cgaaccctag cggagattgc caaagtggag ctggacaata tgccactccg 540  
 tggaaagcag ctgcgtgtgc gctttgcctg ccatagtgc tcccttacag ttcgaaacct 600  
 tcctcagtat gtgtccaacg aactgctgga agaagccttt tctgtgtttg gccaggtaga 660  
 gagggctgta gtcattgttg atgatcgagg aaggccctca ggaaaaggca ttgttgagtt 720  
 ctcagggaag ccagctgctc ggaaagctct ggacagatgc agtgaaggct ccttcctgct 780  
 aaccacattt cctcgtcctg tgactgtgga gcccatggac cagttagatg atgaagaggg 840  
 acttcagag aagctggtta taaaaaacca gcaatttcac aaggaaacgag agcagccacc 900  
 cagatttgca cagcctggct cctttgagta tgaatatgcc atgcgctgga aggcactcat 960  
 tgagatggag aagcagcagc aggaccaagt ggaccgcaac atcaaggagg ctcgtgagaa 1020  
 gctggagatg gagatggaag ctgcacgcca tgagcaccag gtcattgctaa tgagacagga 1080  
 tttgatgagg cgccaagaag aacttcggag gatggaagag ctgcacaacc aagaggtgca 1140

aaaacgaaag caactggagc tcaggcagga ggaagagcgc aggcgccgtg aagaagagat 1200  
 gcggcgccac gaagaagaaa tgatgcggcg acacgaggaa ggattcaagg gaaccttccc 1260  
 tgatgcgaga gagcaggaga ttcggatggg tcagatggct atgggaggtg ctatgggcat 1320  
 aaacaacaga ggtgccatgc cccctgctcc tgtgccagct ggtaccccag ctccctcagg 1380  
 acctgccact atgatgccgg atggaacttt gggattgacc ccaccaacaa ctgaacgctt 1440  
 tggtcaggct gctacaatgg aaggaattgg ggcaattggg ggaactcctc ctgcattcaa 1500  
 ccgtgcagct cctggagctg aatttgcccc aaacaaacgt cgccgatact aataagttgc 1560  
 agtgtctagt ttctcaaaac ctttaaaaga aggacccttt ttggactagc cagaattcta 1620  
 ccctggaaaa gtgttaggga ttctttccaa tagttagatc taccctgcct gtactactct 1680  
 aagggtattc ttccaatagt tagatctacc ctgcctgtac tactctaggg agtatgctgg 1740  
 aggcagaggg caaggagggg gtggtattaa acaatgcaat tctgtgtggg atattgttta 1800  
 atcagttctg tgtggtgcat tcctgaagtc tctaattgtga ctgttgaggg cctggggaaa 1860  
 ccatggcaaa gtggatccag ttagagccca ttaatcttga tcattccggg tttttttttt 1920  
 tttgtccatc ttgtttcatt tgcttgcccc gcccccgaga cggagtctta ctctgtcgcc 1980  
 caggctggag tgtagtgga tgatctcggc tcactgcaat ctctgcctcc cgggttcaag 2040  
 cttgtccagg ttgatcttga actcctgacc tcgtgatcta cccacctcgg tctcccaaaa 2100  
 tgctgggatt acaggggtga gccaccgtgc ccaacctcac ttgcttctta tccttacact 2160  
 cccccagccc cagagaaact gccacataca ccacaaaaac caaacatgcc ccaatgacct 2220  
 tagccccatt gctccattca ctcccagggt agaattcagg caaacgtcca caaaggtcac 2280  
 aggcagcgta catacggttc tggtataccc catatattac cccttcatgt cctaaagaag 2340  
 acattttctc ttagagattt tcattttagt gtatctttaa aaaaaaatc ttgtgttaac 2400  
 ttgcctccat ctttttcttg gggtagggga caccagggaa tgacctttt gtgtctatga 2460  
 tgttgctgtt cacagctttt cttgataggc ctagtacaat cttgggaaca gggttactgt 2520  
 atactgaagg tctgacagta gctcttagac tcgcctatct taggtagtca tgctgtgcat 2580  
 ttttttttct attggtgtac tgtgtttgat ttgtctcata tatttgaggt ttttctgaaa 2640  
 aatggagcag taatgcagca tcaacctatt aaaatacttt taagcctttt 2690

<210> 25

<211> 1828

<212> DNA

<213> Homo sapiens

<400> 25

cagttacagg gagcaccacc agggaaacatc tcggggagcc tggttggaag ctgcaggctt 60  
 agtctgtcgg ctgcgggtct ctgactgccc tgtggggagg gtcttgccctt aacatccctt 120  
 gcatttggct gcaaagaaat ctgcttgga gaaggggtta cgctgtttgg ccgggcagaa 180

actccgctga gcagaacttg ccgccagaat gtcctccttg ttgctgagta tcatcgtcct 240  
 ccacgtcgcg gtgctgggtgc tgctgttcgt ctccacgata gtcagccaat ggatcgtggg 300  
 caatggacac gcaactgatc tctggcagaa ctgtagcacc tcttcctcag gaaatgtcca 360  
 ccactgtttc tcatcatcac caaacgaatg gctgcagtct gtccaggcca ccatgatcct 420  
 gtcgatcatc ttcagcattc tgtctctgtt cctgttcttc tgccaactct tcaccctcac 480  
 caaggggggc aggtttttaca tcaactggaat cttccaaatt cttgctgggtc tgtgcgtgat 540  
 gagtgctgcg gccatctaca cgggtgaggca cccggagtggt catctcaact cggattactc 600  
 ctacggtttc gcctacatcc tggcctgggt ggccttcccc ctggcccttc tcagcgggtgt 660  
 catctatgtg atcttgcgga aacgcgaatg aggcgcccag acggtctgtc tgaggctctg 720  
 agcgtacata gggaaggag gaaggga aaa cagaaagcag acaaagaaaa aagagctagc 780  
 ccaaaatccc aaactcaaac caaaccaaac agaaagcagt ggagggtggg gttgctgttg 840  
 attgaagatg tatataatat ctccggttta taaaacctat ttataaact ttttacatat 900  
 atgtacatag tattgtttgc tttttatgtt gaccatcagc ctctgtttga gccttaaaga 960  
 agtagctaag gaactttaca tcctaacagt ataatccagc tcagtatttt tgttttgttt 1020  
 tttgtttgtt tgttttgtt taccagaaa taagataact ccatctcgcc ccttcccttt 1080  
 catctgaaag aagatacctc cctcccagtc cacctcattt agaaaaccaa agtgtgggta 1140  
 gaaaccccaa atgtccaaaa gcccttttct ggtgggtgac ccagtgcata caacagaaac 1200  
 agccgctgcc cgaacctctg tgtgaagctt tacgcgcaca cggacaaaat gcccaaactg 1260  
 gagcccttgc aaaaacacgg cttgtggcat tggcatactt gcccttacag gtggagtatc 1320  
 ttcgtcacac atctaaatga gaaatcagtg acaacaagtc tttgaaatgg tgctatggat 1380  
 ttaccattcc ttattatcac taatcatcta acaactcac tggaaatcca attaacaatt 1440  
 ttacaacata agatagaatg gagacctgaa taattctgtg taatataaat ggtttataac 1500  
 tgcttttgta cctagctagg ctgctattat tactataatg agtaaatacat aaagccttca 1560  
 tcaactccac atttttctta cggctcggagc atcagaacaa gcgtctagac tccttgggac 1620  
 cgtgagttcc tagagcttgg ctgggtctag gctgttctgt gcctccaagg actgtctggc 1680  
 aatgacttgt attggccacc aactgtagat gtatatatgg tgcccttctg atgctaagac 1740  
 tccagacctt ttgtttttgc tttgcatttt ctgattttat accaactgtg tggactaaga 1800  
 tgcattaaaa taaacatcag agtaactc 1828

<210> 26

<211> 500

<212> DNA

<213> Homo sapiens

<400> 26

gctctcagag gcagcgtgcg ggtgtgctct ttgtgaaatt ccaccatggc gtaccgtggc 60

cagggtcaga aagtgcagaa gggttatgggtg cagcccatca acctcatctt cagatactta 120  
 caaaatagat cgcggattca ggtgtggctc tatgagcaag tgaatatgctg gatagaaggc 180  
 tgtatcattg gttttgatga gtatatgaac cttgtattag atgatgcaga agagattcat 240  
 tctaaaacaa agtcaagaaa acaactgggt cggatcatgc taaaaggaga taatattact 300  
 ctgctacaaa gtgtctccaa ctgaaatga tcaatgaagt gagaaattgt tgagaaggat 360  
 acagtttggt tttagatgtc ctttgcctca tgtgaacatt tattcatatt gttttgatta 420  
 ccctcgtgtt actacaagat ggcaataaat actatgggat tgtttgtatt aaaaaattta 480  
 cattgcttct taaaaaaaaa 500

<210> 27

<211> 4661

<212> DNA

<213> Homo sapiens

<400> 27  
 gctggacttg cctgcggtga cacctgtctc cctctgagag cttcagggtc tccggcctgc 60  
 cttcactggg ttgtgtccag agccggactg attctctcaa tttgcgatct tcagcctgtt 120  
 aaacaagaaa acgaaaaacc cttccagaa aacatggatg catttgaaaa agtgagaaca 180  
 aaattagaaa cacagccaca agaagaatat gaaatcatca atgtggaagt taaacatggg 240  
 ggttttgttt attaccaaga aggttgttgc ttggttcgtt ccaaagatga agaagcagac 300  
 aatgataatt atgaagtttt attcaatttg gaggaactta agttagacca gcccttcatt 360  
 gattgtatca gagttgtctc agatgaaaaa tatgtggctg ccaagataag aactgaagat 420  
 tctgaagcat ctacctgtgt aattataaag ctacgcatc agcccgtaat ggaagcttct 480  
 ttcccgaatg tgtccagttt tgaatgggta aaggacgagg aagatgaaga tgttttattc 540  
 tacaccttcc agaggaacct tcgctgtcat gacgtatatc gagccacttt tgggtgataac 600  
 aaacgtaatg aacgctttta cacagaaaaa gacccaagct actttgtttt cttttatctt 660  
 acaaaagaca gtcgtttcct caccataaat attatgaaca agactacttc tgaagtgtgg 720  
 ttgatagatg gcctgagccc ttgggaccca ccagtactta tccagaagcg aatacatggg 780  
 gtcctttact atgttgaaca cagagatgat gaattataca ttctcactaa tgttggagaa 840  
 cctacagaat ttaagctaag gagaacagcg gctgataccc ctgcaattat gaattgggat 900  
 ttatttttta caatgaagag aaatacaaaa gtgatagact tggacatgtt taaggatcac 960  
 tgtgttctat ttctgaagca cagcaatctc ctttatgtta atgtgattgg tctggctgat 1020  
 gattcagttc ggtctctaaa gctccctcct tgggcctgtg gattcataat ggatacaaat 1080  
 tctgaccaa agaactgccc ctttcaactt tgctctcaa tacgtcccc aaaatattac 1140  
 acatacaagt ttgcagaagg caaactgttt gaggaactg ggcatgaaga cccaatcaca 1200  
 aagactagtc gcgtttttacg tctagaagcc aaaagcaagg atggaaaatt agtgccaatg 1260

actgttttcc acaaaactga ctctgaggac ttgcagaaga aacctctctt ggtacatgta 1320  
 tatggagctt atggaatgga ttgaaaatg aatttcaggc ctgagaggcg ggtcctggtg 1380  
 gatgatggat ggatattagc atactgccat gttcgagggtg gtggtgagtt aggcctccag 1440  
 tggcacgctg atggccgcct aactaaaaaa ctcaatggcc ttgctgattt agaggcttgc 1500  
 attaagacgc ttcattggcca aggccttttct cagccaagtc taacaaccct gactgctttc 1560  
 agtgctggag gggtgcttgc aggagcattg tgtaattcta atccagagct ggtgagagcg 1620  
 gtgacttttg aggcaccttt cttggatggt ctcaacacca tgatggacac tacacttcct 1680  
 ctgacattag aagaattaga agaattgggg aatccttcat ctgatgaaaa acacaagaac 1740  
 tacataaaac gttactgtcc ctatcaaaat attaaacctc agcattatcc ttcaattcac 1800  
 ataacggcat atgaaaacga tgaacgggta cctctgaaag gaattgtaag ttatactgag 1860  
 aaactcaagg aagccatcgc ggagcatgct aaggacacag gtgaaggcta tcagaccctt 1920  
 aatattattc tagatattca gcctggaggc aatcatgtaa ttgaggattc tcacaaaaag 1980  
 attacagccc aaattaaatt cctgtacgag gaacttggac ttgacagcac cagtgttttc 2040  
 gaggatctta agaaatacct gaaattctga aacactgcat tcaactggga attggaaaca 2100  
 cactgaaata tttcatagtc ttacttccaa ttgagttagc aaaaaaaaaa ttaataactt 2160  
 gagactttta agttattaat tttttaaaat gtgcttctcc atctaaattt tgcttagtct 2220  
 acatctcact tgcttatact attcctccat tgatgcacat gccattaac ctaggaaagt 2280  
 agttttcaaa tcatgtcctt tagaaggatg tggagtagag ggaagggaag gattggtgat 2340  
 agcagagctc caggcctccc ttccagtcag aacagttgag cagtttaca attagtgtcc 2400  
 tgcctctttg ctagcaaatg cttttagaca ctgtggcagt gagtcatcct ctaatttcta 2460  
 tgactgcatt ttaagggaaa agataaaatt cttcccctta aaattcgta aagtttttga 2520  
 ataactctgg gtcctaattg gttctggtca tccctgattg atgctatctg aataaagta 2580  
 taagctccta taagccataa tttactttta aacattttat ttttttcaa acatttgaga 2640  
 acctttctta aagcggttac attcaagcta cagaaatata gaagaattaa tgattgttca 2700  
 ccaagcagca tgctgtacat gaagctatta caaatgctta caatccact gaaatgccag 2760  
 tgtcttcac tcttcataaa ggtgcctaac acgagggtata cagtatgttc agtacctgg 2820  
 aatagcatgc tcgattggaa acaaagcatc tatctctgaa agctgtttgg cgatgaagga 2880  
 gattcttcgt gttgtgttca aagatgagtc cctctccctt gtccagaaaa atgccacttg 2940  
 tatcaacttt actgcctttg tcggcagaat tggacttaaa ccttattctt attttagcgg 3000  
 gaaggcccga aatcatatta tgtagattta acagtgttga ttctccaaa ttcagaacca 3060  
 cgataaagat tctgtcagtg ccatccagct ctcttggtga cacaacatag tggctgtcat 3120  
 tcctcaaatt gcaaaaccag cccctgttga ggagtagctc attggcatga agtagactta 3180  
 aatcttgata taacttcaaa gccgatctgg gctgagtcct ttggacctat ttttttaaaa 3240  
 aagtatttac gtaagtgttt gattctaaga attgtttgta agtattttta atatatgtga 3300  
 aggagttatt tacccaaaac acttgctcca attttgcccc ttataattgc caaattgtaa 3360

gcatcaataa gtaggtaaga acaatttata taaaaactga tagaaatgac aaattcgggg 3420  
 tttcggcttg tccgggagtc aataagtacg cacagtgtc tgctacattg tagagtttct 3480  
 gtagagatca aatttgactc cactttagga gtcccaaagc aaatgtccat gtctaagatg 3540  
 aatatttaac ttgcatagtc attctgtgct atattgtaac tgccagatgg ccagaaaagaa 3600  
 ggcaacagtg gactcagact tctgaggaat ttgggtttgt tcccctttgt agactaatgt 3660  
 gtaggttgct gttgtgcgaa gatcgtgtaa ctttagcaga catgtatttc ttgcacagct 3720  
 aatagaagac aaagttgaaa aaaaggatgc aaaataaaaa gctgcctaag gtgaaagtta 3780  
 gaaattgtag actttttttt accataatag tatgtgttca ttgaagatga tttgggttta 3840  
 ttttacagct atataaaaca taatttgatg atgtacttct aacctttcaa gcattttctg 3900  
 ttattgacta tataatatag cctccataaa tgtttttaat gacaatattc tgttgaacgg 3960  
 ttgtaccata ctacagccatg ccttttcatt ttgacgatag tgtttctaatt atttgtatt 4020  
 tttattcccc tccccccatt tttgtattac ttaagataga ttatcagaaa gacagttact 4080  
 ttgtcaaaga gtatgggcac ttgatacata atgccaaatt attcttcata agagctgttg 4140  
 ccaaatacagt gataatgttc atttaattgt attcttgcca gccatgttta ctgggggtgat 4200  
 agttgttatt gtggttggtta ttgttcttta ggggtagggt cccaatatgt ggtcttttaa 4260  
 taattatcta atggtgttta aaaagatgtt tattctgttt gtcagggtaca aagatattta 4320  
 tgatacatgt atgacttgct taagttatta acattttctc tagccttagg taatgcatga 4380  
 aagcacatgt ttcagtgcca ctacataag aagtgcccg taagtgttag ctattattgt 4440  
 ctacttgagt tactactttc taaaagtatg ttgaagtctt tttctgtaat tgcagatttg 4500  
 ttgattttgc atttgagtat tttctatatt ttgaagctgt tagatgcata gtcagatttt 4560  
 ttggtggaat gttttatcaa tttttgaaaa ttgcctttgt ctcatataat gcttttcata 4620  
 ttgaactata tttgtctgc tattaaatac ttccaagcct g 4661

<210> 28

<211> 1135

<212> DNA

<213> Homo sapiens

<400> 28

ggatccggca acgaaggtac catggccgga ctccggagcc gcacaaacca gggctcgcca 60  
 tgaagccagg attcagtcct cgtgggggtg gctttggcgg ccgagggggc tttggtgacc 120  
 gtggtggtcg tggaggccga gggggctttg gcgggggccc aggtcgaggc ggaggcttta 180  
 gaggtcgtgg acgaggagga ggtggaggcg gcggcgccgg tggaggagga ggaagagggtg 240  
 gtggaggcct ccatctgtgt ggcaaccggg gtcgtggctg gggaggaaaa agaggaaacc 300  
 agtcggggaa gaatgtgatg gtggagccgc atcggcatga ggggtgtctt atttgtcgag 360  
 gaaaggaaga tgcactggtc accaagaacc tgggtccctg ggaatcagtt tatggagaga 420

agagagtctc gatttcggaa ggagatgaca aaattgagta ccgagcctgg aacccttcc	480
gctccaagct agcagcagca atcctgggtg gtgtggacca gatccacatc aaaccggggg	540
ctaaggttct ctacctcggg gctgcctcgg gcaccacggt ctcccatgtc tctgacatcg	600
ttggtccgga tgggtctagtc tatgcagtcg agttctccca ccgctctggc cgtgacctca	660
ttaacttggc caagaagagg accaacaatca ttcctgtgat cgaggatgct cgacaccac	720
acaaataaccg catgctcatc gcaatggtgg atgtgatctt tgctgatgtg gcccagccag	780
accagacccg gattgtggcc ctgaatgccc acaccttcct gcgtaatgga ggacactttg	840
tgatttccat taaggccaac tgcattgact ccacagcctc agccgaggcc gtgtttgcct	900
ccgaagtga aaagatgcaa caggagaaca tgaagccgca ggagcagttg acccttgagc	960
catatgaaag agaccatgcc gtggtcgtgg gagtgtacag gccaccccc aaggtgaaga	1020
actgaagttc agcgtgtca ggattgagag agatgtgtgt tgatactgtt gcacgtgtgt	1080
ttttctatta aaagactcat ccgtcaaaaa aaaaaaaaaa aaaaaaaa aaaaa	1135

&lt;210&gt; 29

&lt;211&gt; 6734

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<400> 29	
cccagttgtc tgcgggctgc ggggagctaa gtccccagat tggaggaggc tggctctggt	60
cttcgatgca caggagtggc cgttatggaa gcgagcagca gcgtgcaggg tcaaagacag	120
ccggccccc atgtcagtg tctaggatgg ccagtgaagg caccaacatc ccaagtcctg	180
tgggtgcgca gattgacaag cagtttctga tttgcagtat atgcctggaa cgggtacaaga	240
atcccaaggt tctcccctgt ctgcacactt tctgcgagag gtgcctgcag aactacattc	300
ctgcccacag tttaacctc tcctgcccag tgtgccgcca gacctcatc ctgcccgaga	360
aaggggtggc cgcgctccag aacaatttct tcatcacaaa cctgatggac gtgctgcagc	420
gaactccagg cagcaacgct gaggagtctt ccatacctgga gacagtcact gctgtggctg	480
cgggaaagcc tctctcttgc ccaaaccacg atgggaatgt gatggaattt tactgccagt	540
cctgtgagac tgccatgtgt cgggagtgca cggaggggga gcacgcagag caccacacag	600
ttccactcaa ggatgtggtg gaacagcaca aggcctcgt ccagggtccag ctggatgctg	660
tcaacaaaag gctcccagaa atagattctg ctcttcagtt catctctgaa atcattcatc	720
agttaaccaa ccaaaaggcc agcatcgtgg atgacattca ttccacctt gatgagctcc	780
agaagacttt aaatgtgcgc aagagtgtgc tgcttatgga attggaggtc aactatggcc	840
tcaaacacaa agtctccag tcgcagctgg atactctgct ccaggggcag gagagcatta	900
agagctgcag caacttcaca gcgcaggccc tcaacatgg cacggagacc gaggtcctac	960
tgggtgaagaa gcagatgagc gagaagctga acgagctggc cgaccaggac ttccccttgc	1020



acccgcgga gaacgaccag ctggatttca tcgtggaaac cgaggggctg aagaagtcca 1080  
 tccacaacct cgggacgatc ttaaccacca acgccgttgc ctcagagaca gtggccacgg 1140  
 gcgaggggct gcggcagacc atcatcgggc agcccatgtc cgtcaccatc accaccaagg 1200  
 acaaagacgg tgagctgtgc aaaaccggca acgcctacct caccgccgaa ctgagcacc 1260  
 ccgacgggag cgtggcagac ggggagatcc tggacaacaa gaacggcacc tatgagtttt 1320  
 tgtacactgt ccagaaggaa ggggacttta ccctgtctct gagactctat gaccagcaca 1380  
 tccgaggcag cccgttttaag ctgaaagtga tccgatccgc tgatgtgtct cccaccacag 1440  
 aaggcgtgaa gaggcgcgtt aagtccccgg ggagcggcca cgtcaagcag aaagctgtga 1500  
 aaagaccgc aagcatgtac agcactggaa aacgaaaaga gaatcccatc gaagacgatt 1560  
 tgatctttcg agtgggtacc aaaggaagaa ataaaggaga gtttacaat cttcagggg 1620  
 tagctgcatc tacaatgga aagatattaa ttgcagacag taacaaccaa tgtgtgcaga 1680  
 tattttccaa tgatggccag ttcaaaagtc gttttggcat acggggacgc tctccggggc 1740  
 agctgcagcg gccacagga gtggctgtac atcccagtg ggacataatc attgccgatt 1800  
 atgataataa atgggtcagc attttctcct ccgatgggaa atttaagaca aaaattggat 1860  
 caggaaagct gatgggaccc aaaggagttt ctgtggaccg caatgggcac attattgttg 1920  
 tggacaacaa ggcgtgtgc gtgtttatct tccagccaaa cgggaaaata gtcaccaggt 1980  
 ttggtagccg aggaaatggg gacaggcagt ttgcagggtcc ccattttgca gctgtaaata 2040  
 gcaataatga gattattatt acagatttcc ataatcattc tgtcaagggtg tttaatcagg 2100  
 aaggagaatt catgttgaag tttggctcaa atggagaagg aaatgggcag tttaatgctc 2160  
 caacagggtg agcagtggat tcaaatggaa acatcattgt ggccgactgg ggaaacagca 2220  
 ggatccagggt ttttgatggg agtggatcat ttttgccta cattaacaca tctgctgacc 2280  
 cactctatgg cccccaaggc ctggccctaa cttcagatgg tcatgttgtg gttgcagact 2340  
 ctggaaatca ctgtttcaaa gtctatcgat acttacagta atggtgggca ggtggatacc 2400  
 cgcttccatg gtcttgact ataaactgga atggatttct caatgcggga ccagattatg 2460  
 actagagttt ttatgccaga aggaatcatt ggtgaacttt ccaaggttat ttctgaatgt 2520  
 aacaatttcc ttaaaaatga cttatccaat ttctgtattt caccttttagg gttaaaaaa 2580  
 actcttctac tgaatctata aaaactgcag ttttacatct gtgaactatg gcttaaggga 2640  
 caggatttat gtagctaaac taattttgca aatcaaacag aacttaaaa aaactagcat 2700  
 atgtaaagggt attcgtaaat cctgtgaatg gtagcttttg cacagaactt ccaaagcaa 2760  
 aacaaaaaca aaatctattg tagttatata cttcatttaa ctaggtcac aagaccagg 2820  
 gaatcttcta acctcacttt tacagtaggt attactcttg tgacattttt ttggttatca 2880  
 acaactaaat ataaattact ttggaaaaag taaggctgtc ttgcaaatg atcccagctc 2940  
 tgattagcag ccctctggag ttcagaactt aagtatcagt gcaaatttct caacctttct 3000  
 gggttagaca aagatccttt tttgtgtgtt cttttcacca cccctttggc tcaccttgta 3060  
 tcagcaaaca aagtacttct tcagggaac ctgaaatttc taatgccttg aaaagcatat 3120

tacaaaagta atgctacctt ttgggaaaca aactgccccg ttaactccag atcattgcac	3180
tggaatgtaa tcaagaaagt tagtcatgtt ttatgtacca tgttttcaca cgtgtctctt	3240
ctcttcgact tcctgaaagc gaaagcttta cctcctgcaa atgtcagcac atgtagtagg	3300
acaccagtat ctaggacag agagccataa gtagcccttt ggaggactga tgggtgtcaac	3360
caaaggcatg tgattgatta atgattcccc cttagaaagc aagtgttacc aaagttgtgt	3420
tatcttgaaa gcattacagg taagggcatg ttatggttat ttatcattgt ttaatgaata	3480
gtagaggtgt caagggacta tgtatacatg attagggtaa gatagaatgt attatatata	3540
tatatatata tacacacaca catatatata gctgaatctt tgggtgtattg aaataggcag	3600
cactctgaaa gacagaagct tcgtccagcc actcttcagc acattccttt actaagcagt	3660
ttaaagccgt ctagtggag caagccctaa agcagattta atttttgcca ttttccaaga	3720
atgacggtgg tggcttttag tcagaaaatg gccttctgtg ctttcaaaaa aaaaacaaaa	3780
aaaaaaccac acacacacat aaaaaacca acaggtcaaa ataaaagttg aacttgagtt	3840
acatttaatt taaatataaa tgcattttga gaaatgttaa gaacaattta gtcaatcgtt	3900
catctgtcat tggtagtga aaataagctg tgggtctattt ccactgttta attttctact	3960
cagttctacc aaataggatg tcatgtttga ctttttgat agtgactttg ggggtcttctt	4020
cactgaaagc accttagaac tgtactataa gaaaacattt cccctatgta taattatatg	4080
aatgtgatgt ttattgctta ttaatttata attcagtcatt tctctatata ggacttctta	4140
aaatttagaa gggaaatcta gctacttcaa attgtctgtt aaatttatta tgcccaaatac	4200
aacctctgaa aaaaggtttt tccaggaaga ttacattta ggtttaatat ttttttagtt	4260
aggtagagtt ttaaaaaata cttgagcctg tccgtgataa agctataaaa ttcaataact	4320
ttttagaatg ttaaatgaag aactgtttc ctaacatcag tgagatacat ctttgaattt	4380
aaacattcat atttactgag tacctactag gtaccaagta ctcttttagg cactggaaat	4440
acagtgatgg acaaaacagg taaaaaatcg ctgccccctc agagctgaca ttctggggtg	4500
ggaatttcat ttgcccagct actaacgttc tgcacaaaag acaggctaga ctcttgtcta	4560
gattgtttaa aagaaacttt tcaaattgggt tacattaatt ttagtttatt ttcacaagta	4620
aaaatggctt ttattttaga ttctttctgt cccaggctgt tgatcttaaa actagttgat	4680
ttaaagagtt tttttgcaca acatttcaat tatatttggt aacttagaaa ttaacttaca	4740
atctaaccag ccatcatatc atatcctatc aggctagata tctcaatagt agactgaata	4800
caaagctaat tttttttaca tgtcaatatt ggcacaaact ggaatgaaag aatagtttga	4860
ttcagacctg ctccactatg tgttgctaaa acacatgcta tgagcactcg aggaaacact	4920
atattttttc caaaaaatat gtgattatat atgttaaagt atagataaca tttcacactt	4980
ggatacatat gtgcatttac tgtatttctt ggtaagcata tttttggggg aaagtgtctgc	5040
tgatatgata caagtagaca aaattttaat gaaattttgt cacattctat ggaaatgggt	5100
ttctggtaaa ctgagaagga tattaataa agtggtcttt ttctgggcta ccattattgt	5160
ttgatttctc ttgtcaagt gtatagaacc tgtcatatc tcatgataag tagcactgaa	5220

aaattactca ttcaaatttc ccctgggcac gtaaggcaaa atattgccgg ttgggatttc 5280  
 aaggtcagtg acgacgcatt tcctcccagt acagaccccc cagccccctt tgctggacat 5340  
 ggggaggcag agagtcactt gaccatccag aaatacatga ctacaagtcc tttatgactg 5400  
 ttgcccattt tttttaatgg tacttagtat ttgatcaaa ctttagtctc cagaactaaa 5460  
 caagtcacct agtttcctta ttttaattta ctgtgactag atttgaagca aataaatact 5520  
 ccagatccat gcagctagaa cacacttgct tccactacta aatatacagg gtatgtccta 5580  
 acatggagtt aactggaata gcagtacact agcaagtatc tgtgaatcct tagcactgac 5640  
 ggggttaacag aaatgctttg gtaataccta cttagttaat tggaggaagt agtaaataaa 5700  
 cattaggtaa tctgcagatt acttcaaagt ggaaaaatct tttgtagac tctatagtac 5760  
 cctctctatt cactagcttc tgaaaagggg ggagtatttt tagtttgaca atttaataat 5820  
 ttaaaaaaaa gacatctcca ggtaggaaaa aatgaaagct atttcatgca aacattatct 5880  
 aatttagctt aaaagtgaag gtggaataac tgttggtttc tgtaaagtgt gcagggtttt 5940  
 aaactttata attactttaa tttttttgat aactagaaat ctagtattgc cataaaggaa 6000  
 actaagtgcc catcaaagat ttgttttgta taaataaaga attatttggt ttgttttcaa 6060  
 tgacagtaag ctacaaatca tgatgcttaa aaactttcta aagatgaatt gtgtggcagt 6120  
 gattggctctg tttgtggaga atgtatgaaa gctattaata ttctagaata gattaataaa 6180  
 ttggctatgt tgttccaatg aatgtacagc acttccatta acttttgaaa gcaacacagc 6240  
 cttaaaactca atgcttttgc tttatgacat gggaatgttc tgtcatcaat ggagtgtatt 6300  
 cttgtaatag aattctttat atcgttctca attctataga ctttcaagcc tatgtatgaa 6360  
 tatgaagggg tttttttttt tttgctttgt tttcttttta gattttgtac attccatctt 6420  
 tataggctctg tttcatatgt tttatgtata gaacactaag tcttgactc tctgacattg 6480  
 atactgatat attctcgtca tttgttcttt tatgaatcaa aatgttgact gcctatttaa 6540  
 agaaaagaat gaacgctgtg catcaaagtg tttgtatgtt cgtagctaca tacgtaccac 6600  
 agtatttttg atgcttttag ctacaatgaa actttcaatt aattctgtct tgaaacatag 6660  
 gagaaacagg attcatgtgt atctctttac catgcacaaa atctcaaatc attataataa 6720  
 agcttggttt ctcc 6734

<210> 30

<211> 3744

<212> DNA

<213> Homo sapiens

<400> 30

ccacgcgtcc ggtggcggtc gagcgtggcg taggcgaatc ctcggcacta agcatatgga 60  
 cctcgcggcg gcagcggagc cgggcgccgg cagccagcac ctggagggtc gcgacgaggt 120  
 ggccgagaag tgccagaaac tgttcctgga cttcttgag gagtttcaga gcagcgatgg 180

agaaattaaa	tacttgcaat	tagcagagga	actgattcgt	cctgagagaa	acacattggt	240
tgtgagtttt	gtggacctgg	aacaatttaa	ccagcaactt	tccaccacca	ttcaagagga	300
gttctataga	gtttaccctt	acctgtgtcg	ggccttgaaa	acattcgtca	aagaccgtaa	360
agagatccct	cttgccaagg	atttttatgt	tgcattccaa	gacctgccta	ccagacacaa	420
gattcgagag	ctcacctcat	ccagaattgg	tttgctcact	cgcatcagtg	ggcagggtgg	480
gcggaactcac	ccagttcacc	cagagcttgt	gagcggaact	tttctgtgct	tggactgtca	540
gacagtgatc	agggatgtag	aacagcagtt	caaatacaca	cagccaaaca	tctgccgaaa	600
tccagtttgt	gccaacagga	ggagattcct	actggatata	aataaatcaa	gatttggtga	660
ttttcaaaaag	gttcgtattc	aagagaccca	agctgagctt	cctcgaggga	gtatcccccg	720
cagtttagaa	gtaattttta	gggctgaagc	tgtggaatca	gctcaagctg	gtgacaagtg	780
tgactttaca	gggacactga	ttgttggtgc	tgacgtctcc	aagcttagca	caccaggagc	840
acgtgcagaa	actaattccc	gtgtcagtg	tggtgatgga	tatgagacag	aaggcattcg	900
aggactccgg	gcccttggtg	ttagggacct	ttcttatagg	ctggctcttc	ttgcctgctg	960
tgttgcgcca	accaacccaa	ggtttggggg	gaaagagctc	agagatgagg	aacagacagc	1020
tgagagcatt	aagaacccaa	tgactgtgaa	agaatgggag	aaagtgtttg	agatgagtca	1080
agataaaaat	ctataccaca	atctttgtac	cagcctgttc	cctactatac	atggcaatga	1140
tgaagtaaaa	cggggtgtcc	tgctgatgct	ctttggtggc	gttccaaaga	caacaggaga	1200
agggacctct	cttcgagggg	acataaatgt	ttgcattggt	ggtgacccaa	gtacagctaa	1260
gagccaattt	ctcaagcacg	tggaggagtt	cagccccaga	gctgtctaca	ccagtggtaa	1320
agcgtccagt	gctgctggct	taacagcagc	tggtgtgaga	gatgaagaat	ctcatgagtt	1380
tgctattgag	gctggagctt	tgatgttggc	tgataatggg	gtgtgttgta	ttgatgaatt	1440
tgataagatg	gacgtgcggg	atcaagttgc	tattcatgaa	gctatggaac	agcagaccat	1500
atccatcact	aaagcaggag	tgaaggctac	tctgaacgcc	cggacgtcca	ttttggcagc	1560
agcaaaccga	atcagtggac	actatgacag	atcaaaatca	ttgaaacaga	atataaattt	1620
gtcagctccc	atcatgtccc	gattcgatct	cttctttatc	cttggtggatg	aatgtaatga	1680
ggttacagat	tatgccattg	ccaggcgcat	agtagatttg	cattcaagaa	ttgaggaatc	1740
aattgatcgt	gtctattccc	tcgatgatat	cagaagatat	cttctctttg	caagacagtt	1800
taaacccaag	atttccaaag	agtcagagga	cttcattgtg	gagcaatata	aacatctccg	1860
ccagagagat	ggttctggag	tgaccaagtc	ttcatggagg	attacagtgc	gacagcttga	1920
gagcatgatt	cgtctctctg	aagctatggc	tcggatgcac	tgctgtgatg	aggtccaacc	1980
taaacatgtg	aaggaagctt	tccggttact	gaataaatca	atcatccgtg	tggaaacacc	2040
tgatgtcaat	ctagatcaag	aggaagagat	ccagatggag	gtagatgagg	gtgccgggtg	2100
catcaatggg	catgctgaca	gccctgctcc	tgtgaacggg	atcaatggct	acaatgaaga	2160
cataaatcaa	gagtctgctc	ccaaagcctc	cttaaggctg	ggcttctctg	agtactgccg	2220
aatctctaac	cttattgtgc	ttcacctcag	aaagggtggaa	gaagaagagg	acgagtcagc	2280

attaaagagg agcgagcttg ttaactggta cttgaaggaa atcgaatcag agatagactc 2340  
 tgaagaagaa cttataaata aaaaaagaat catagagaaa gttattcatc gactcacaca 2400  
 ctatgatcat gttctaattg agctcaccca ggctggattg aaaggctcca cagagggaaag 2460  
 tgagagctat gaagaagatc cctacttggg agttaaccct aactacttgc tcgaagattg 2520  
 agatagtga agtaactgac cagagctgag gaactgtggc acagcacctc gtggcctgga 2580  
 gcctggctgg agctctgcta gggacagaag tgtttctgga agtgatgctt ccaggatttg 2640  
 ttttcagaaa caagaattga gttgatggtc ctatgtgtca cattcatcac aggtttcata 2700  
 ccaacacagg cttcagcact tcctttgggtg tgtttctgtg ccagtgaaag ttggaaccaa 2760  
 ataatgtgta gtctctataa ccaatacctt tgtttctatg tgtaagaaaa ggcccattac 2820  
 ttttaaggta tgtgctgtcc tattgagcaa ataacttttt ttcaattgcc agctactgct 2880  
 tttattcatc aaaataaaat aacttggttct gaagttgtct attggatttc tttctactgt 2940  
 accctgatta ttacttccat ctacttctga atgtgagact ttcccttttt gcttaacctg 3000  
 gagtgaagag gtagaactgt ggtattatgg atgaggtttc tatgagaagg agtcattaga 3060  
 gaactcatat gaaagctaga ggccttagag atgactttcc aagggttaatt ccagtttttt 3120  
 ttttttttaa gtttataaaa gtttattata ctttttttaa attactcttt agtaatttat 3180  
 tttacttctg tgtcctaagg gtaatttctc aggattgttt tcaaattgct tttttagggg 3240  
 aaataggcca tttgctatat tacaagcaat ccccaaattt tatggtcttc caggaaaagt 3300  
 tattaccgtt tatgatacta acagttcctg agacttagct atgatcagta tgttcatgag 3360  
 gtggagcagt tcctgtgttg cagcttttaa caacagatgg cattcattaa atcacaaagt 3420  
 atgttaaagg tcacaaaagc aaaataactg tctgaggcta aggcccacgt gggacagtct 3480  
 aatacccatg agtactcaac ttgccttgat gtctgagctt tccagtcaa tgtgaatttg 3540  
 agcagccaga aatctattag tagaaagcaa gacagattaa tataggtaa aacaatgatt 3600  
 taaatatgtt tctccaata attatctctt tccctggaat caacttgat gaaaccttgt 3660  
 caaatgtac tccacaagta tgtacaatta agtattttta aaataaatgg caaacattaa 3720  
 aaaaaaaaaa aaaaaaaaaa aaaa 3744

<210> 31

<211> 3321

<212> DNA

<213> Homo sapiens

<400> 31

ttgtgagtct ataactcgga gccgttgggt cggttctgc tattccggcg cctccactcc 60  
 gtcccccgcg ggtctgctct gtgtgccatg gacggcattg tcccagatat agccgttggg 120  
 acaaagcggg gatctgacga gcttttctct acttgtgtca ctaacggacc gtttatcatg 180  
 agcagcaact cggcttctgc agcaaacgga aatgacagca agaagttaa aggtgacagc 240

cgaagtgcag	gcgtcccctc	tagagtgatc	cacatccgga	agctccccat	cgacgtcacg	300
gagggggaag	tcattctcct	ggggctgccc	tttggaagg	tcaccaacct	cctgatgctg	360
aaggggaaaa	accaggcctt	catcgagatg	aacacggagg	aggctgcaa	caccatggtg	420
aactactaca	cctcggtgac	ccctgtgctg	cgcgccagc	ccatctacat	ccagttctcc	480
aaccacaagg	agctgaagac	cgacagctct	cccaaccagg	cgcgggcca	ggcgccctg	540
caggcggtga	actcggtcca	gtcggggaac	ctggccttgg	ctgcctcggc	ggcgccctg	600
gacgcaggga	tggcgatggc	cgggcagagc	cccgtgctca	ggatcatcgt	ggagaacctc	660
ttctaccctg	tgaccctgga	tgtgctgcac	cagattttct	ccaagttcgg	cacagtgttg	720
aagatcatca	ccttcaccaa	gaacaaccag	ttccaggccc	tgctgcagta	tgcggaaccc	780
gtgagcgccc	agcacgcaa	gctgtcgctg	gacgggcaga	acatctacaa	cgctgctgc	840
acgctgcgca	tcgacttttc	caagctcacc	agcctcaacg	tcaagtacaa	caatgacaag	900
agccgtgact	acacacgccc	agacctgcct	tccggggaca	gccagccctc	gctggaccag	960
accatggccg	cggccttcgg	tgcacctggt	ataatctcag	cctctccgta	tgcaggagct	1020
ggtttcctc	ccacctttgc	cattcctcaa	gctgcaggcc	tttccgttcc	gaacgtccac	1080
ggcgccctg	cccccttggc	catcccctcg	gcggcgccg	cagctgcggc	ggcaggctcg	1140
atcgccatcc	cgggcctggc	gggggcagga	aattctgtat	tgctggtcag	caacctcaac	1200
ccagagagag	tcacacccca	aagcctcttt	attcttttcg	gcgtctacgg	tgacgtgcag	1260
cgctgaaga	tcctgttcaa	taagaaggag	aacgccctag	tgcagatggc	ggacggcaac	1320
cagggccagc	tggccatgag	ccacctgaac	gggcacaagc	tgacgggaa	gcccacccgc	1380
atcacgctct	cgaagcacca	gaacgtgcag	ctgccccg	agggccagga	ggaccagggc	1440
ctgaccaagg	actacggcaa	ctcaccctg	caccgcttca	agaagccggg	ctccaagaac	1500
ttccagaaca	tattcccgc	ctcggccacg	ctgcacctct	ccaacatccc	gccctcagtc	1560
tccgaggagg	atctcaaggt	cctgttttcc	agcaatggg	gcgtcgtcaa	aggattcaag	1620
ttcttcaga	aggaccgcaa	gatggcactg	atccagatgg	gctccgtgga	ggaggcggtc	1680
cagggccctca	ttgacctgca	caaccacgac	ctcggggaga	accaccacct	gcgggtctcc	1740
ttctccaagt	ccaccatcta	ggggcacagg	ccccacggc	cgggccccct	ggcgacaact	1800
tccatcattc	cagagaaaag	ccactttaaa	aacagctgaa	gtgaccttag	cagaccagag	1860
atcttatctt	tttaaagaga	aatcagttta	cctgttttta	aaaaaattaa	atctagtcca	1920
ccttgctcac	cctgcggtga	cagggacagc	tcaggctctt	ggtgactgtg	gcagcgggag	1980
ttcccggccc	tccacacccg	gggacagacc	ctcggggcca	tgcttggtg	gggcctgtgt	2040
cgggcgtggg	gcctgcaggt	gggcgccccg	accacgactt	ggcttccttg	tgcttaaaa	2100
aacctgcctt	cctgcagcca	cacaccaccc	cggggtgtcc	tggggacca	aggggtgggg	2160
gggtcacacc	agagagaggc	agggggcctg	gccggctcct	gcaggatcat	gcagctgggg	2220
cgcgcgggcc	gcggctgcga	cacccaacc	ccagccctct	aatcaagtca	cgtgattctc	2280
ccttcacccc	gccccagg	ccttccttc	tgccccagg	cgggctcccc	gctgctccag	2340

ctgCGGagct ggtcgacata atctctgtat tatatacttt gcagttgcag acgtctgtgc 2400  
 ctagcaatat ttccagttga ccaaataattc taatcttttt tcatttatat gcaaaagaaa 2460  
 tagttttaag taacttttta tagcaagatg atacaatggg atgagtgtaa tctaaacttc 2520  
 cttgtgggtat taccttgtat gctgttactt ttattttatt ccttghtaatt aagtcacagg 2580  
 caggacccag tttccagaga gcaggcgggg ccgcccagtg ggtcaggcac agggagcccc 2640  
 ggtcctatct tagagcccct gagcttcagg gaaggggagg gcgtgtcgcc gcctctggca 2700  
 tcgcctccgg ttgccttaca ccacgccttc acctgcagtc gcctagaaaa cttgctctca 2760  
 aacttcaggg ttttttcttc cttcaaattt tggaccaaag tctcatttct gtgttttgcc 2820  
 tgcctctgat gctgggaccc ggaaggcggg cgctcctcct gtcttctctg tgctctttct 2880  
 accgcccccg cgtcctgtcc cgggggctct ctaggatcc ctttccgta aaagcgtgta 2940  
 acaaggggtg aaatatattt aattttttat acctgttggt agaccgagg ggcgggcgcg 3000  
 cggtttttta tggtgacaca aatgtatatt ttgctaacag caattccagg ctcagtattg 3060  
 tgaccgcgga gccacagggg accccacgca cattccgttg ccttaccga tggcttgtga 3120  
 cgcgagaga accgattaaa accgtttgag aaactcctcc cttgtctagc cctgtgttcg 3180  
 ctgtggacgc tgtagaggca ggttggccag tctgtacctg gacttcgaat aaatcttctg 3240  
 tatcctcgct ccgttccgcc ttaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3300  
 aaaaaaaaaa aaaaaaaaaa a 3321

<210> 32

<211> 1209

<212> DNA

<213> Homo sapiens

<400> 32

gaattcctga cttccttttc ggaggaagat ccttgagcag ccgacgttgg gacaaaggat 60  
 ttggagaaac ccagggctaa agtcacgttt ttctccttt aagacttacc tcaacacttc 120  
 actccatggc agtccccgag acccgcccta accacactat ttatatcaac aacctcaatg 180  
 agaagatcaa gaaggatgag ctaaaaaagt ccctgtacgc catcttctcc cagtttggcc 240  
 agatcctgga tatcctggta tcacggagcc tgaagatgag gggccaggcc tttgtcatct 300  
 tcaaggaggt cagcagcgcc accaacgccc tgcgtccat gcagggtttc cttttctatg 360  
 acaaacctat gcgtatccag tatgccaaga ccgactcaga tatcattgcc aagatgaaag 420  
 gcaccttcgt ggagcgggac cgcaagcggg agaagaggaa gcccagagc caggagaccc 480  
 cggccaccaa gaaggctgtg caaggcgggg gagccacccc cgtggtgggg gctgtccagg 540  
 ggctgtccc gggcatgccg ccgatgactc aggcgccccg cattatgcac cacatgccgg 600  
 gccagccgcc ctacatgccg ccccttggtg tgatccccc gccaggcctt gcacctggcc 660  
 agatcccacc aggggccatg ccccgccagc agcttatgcc aggacagatg cccctgccc 720

```

agcctctttc tgagaatcca ccgaatcaca tcttgttcct caccaacctg ccagaggaga 780
ccaacgagct catgctgtcc atgcttttca atcagttccc tggcttcaag gaggtccgtc 840
tggtaccgag gcggcatgac atcgccctcg tggagtttga caatgaggta caggcagggg 900
cagctcgaga tgccttcgag ggctttaaga tcacgcagaa caacgccatg aagatctcct 960
ttgccaagaa gtagcacctt ttccccccat gcctgcccct tcccctgttc tggggccacc 1020
cctttccccc ttggctcagc cccctgaagg taagtccccc cttggggggc ttcttgagc 1080
cgtgtgtgag tgagtggtcg ccacacagca ttgtaccag agtctgtccc cagacattgc 1140
acctggcgct gttaggccgg aattaaagtg gctttttgag gtttggtttt tcacaaaaaa 1200
aaggaattc 1209

```

&lt;210&gt; 33

&lt;211&gt; 1432

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

```

<400> 33
gctgttcggc ctgctgcgt ccgggagctg ccgacggacg gagcgcccc gcccccgcc 60
ggccgccgc ccgccgcgc catgccctt tccaacagcc acaacgact gaagctgcgc 120
ttcccgccg aggacgagtt ccccgacct agcgcccaca acaaccacat ggccaagggtg 180
ctgaccccc agctgtacgc ggagctgcgc gccaaagca cgccgagcgg cttcacgctg 240
gacgacgtca tccagacagg cgtggacaac ccggggccacc cgtacatcat gaccgtgggc 300
tgctggcgg gcgacgagga gtcctacgaa gtgttcaagg atctcttga ccccatcatc 360
gaggaccggc acggcggtca caagcccagc gatgagcaca agaccgacct caaccccgac 420
aacctgcagg gcggcgacga cctggacccc aactacgtgc tgagctcgc ggtgcgcacg 480
ggccgcagca tccgtggctt ctgcctcccc ccgactgca gccgcgggga gcgccgcgcc 540
atcgagaagc tcgcggtgga agccctgtcc agcctggacg gcgacctggc gggccgatac 600
tacgcgtca agagcatgac ggaggcggag cagcagcagc tcatcgacga ccacttcctc 660
ttcgacaagc ccgtgtcgcc cctgtgtgtg gcctcgggca tggcccgca ctggcccgac 720
gcccgcggta tctggcaca tgacaataag accttcctgg tgtgggtcaa cgaggaggac 780
cacctgcggg tcattccat gcagaagggg ggcaacatga aggaggtgtt caccgccttc 840
tgaccggcc tcaccagat tgaaactctc ttcaagtcta aggactatga gttcatgttg 900
aacctcacc tgggtacat cctcacctgc ccatccaacc tgggcaccgg gctgcgggca 960
ggtgtgcata tcaagctgcc caacctgggc aagcatgaga agttctcga ggtgcttaag 1020
cggctgcgac ttcagaagcg aggcacagcg ggtgtggaca cggctgcggg gggcggggtc 1080
ttcgacgtct ccaacgctga ccgcctgggc ttctcagagg tggagctggg gcagatgggtg 1140
gtggacggag tgaagctgct catcgagatg gagcagcggc tggagcaggg ccaggccatc 1200

```



gacgacctca tgcctgcccc gaaatgaagc ccggtccaca cccgacacca gccctgctgc 1260  
 ttcctaactt attgcctggg cagtgtccac catgcacccc tgatgttcgc cgtctggcga 1320  
 gcccttagcc ttgctgtaga gacttccgtc acccttggtg gagtttattt ttttgatggc 1380  
 taagatactg ctgatgctga aataaactag ggttttggtc tgcctgcgtc tg 1432

<210> 34

<211> 3309

<212> DNA

<213> Homo sapiens

<400> 34

gcggcgcgcc cgagcctagt cccacgccc cggcgcgccc gggctccctg ctgatcccag 60  
 aacaatcaac catgacgacc gaatctggat cagactcggg atccaagccg gaccaggagg 120  
 ccgagcccca ggaggcggcg ggggcgcagg ggcggcgggg gccgtgccgg agccgcccga 180  
 ggaggagcag cagcaggccc tggagcagtt cggcgccgct gcagcgcaca gcaccccggt 240  
 gcgaggggagg tcaactgaca ggaacaggag ttgctgcca gggctgcaaa acagctcgaa 300  
 tatcagcaat tagaagacga taaactttct cagaaatcat ctagcagtaa actctctcgg 360  
 tctccattaa agattgtcaa aaagcctaaa agcatgcagt gcaaagtgat acttctcgat 420  
 ggatcagaat atacctgtga tgtagagaaa cgctccagag gacaagtgct gtttgataaa 480  
 gtgtgtgaac acttgaactt gctagagaaa gactactttg ggcttacgta tcgagatgct 540  
 gaaaaccaga agaattgggt ggaccctgct aaggaaataa aaaaacagggt tcgaagtggg 600  
 gcttggcact tttcatttaa tgtgaaattt tatccaccag accctgccc actatctgaa 660  
 gatatacaca ggtactacct ctgcttgcat ttgcgagatg acatcgtgtc cggaaggctg 720  
 ccctgctcct ttgttaccct ggccttgctg ggctcctaca ctgtccagtc agagctcggg 780  
 gactatgacc cagatgaatg tgggagcgtg tacattagtg agttccgctt tgcaccaaac 840  
 cacactaaag aactggaaga caaagtgatc gagctgcaca agagccacag aggaatgacg 900  
 ccagcagaag cagagatgca tttcttgtaa aatgccaaaa aattatcaat gtatggggta 960  
 gatttacatc atgctaagga ctcagaaggg gtagaaatta tgtaggagt ttgtgcaagt 1020  
 ggtctgttga tatatcgca ccggctgcga ataaacagat ttgcctggcc caaggttcta 1080  
 aagatttcat acaaacggaa caacttttac attaatgacc ggccgggaga gtttgaacaa 1140  
 tttgaaagca ccattgggtt taagctgcca aaccatcgag ctgccaagcg tttatggaaa 1200  
 gtatgtgttg agcatcatal atttttcaga ctactgttac cagaagcacc tcccaagaaa 1260  
 ttcctaacct tgggttccaa gtttcgttat agtggcagga cacaagcgca aacgagaaga 1320  
 gccagtgcgt tgatagatcg ccagcacct tactttgaac gctcatccag caaacgttat 1380  
 accatgtctc gcagcttgga tggagcatca gtgaatgaaa accatgaaat atacatgaag 1440  
 gattctatgt ctgctgcaga ggttggtact ggccagtacg ccacaacaaa aggcattctc 1500

cagaccaact tgatcaccac tgtgactccg gagaagaagg ctgaggagga gcgggacgag 1560  
 gaagaggaca aacggaggaa gggggaagaa gtcacgcca tctcggccat ccagcacgag 1620  
 ggaaagactg acagtgagcg cacggacacc gcagccgacg gggagaccac tgccactgag 1680  
 gagctagaaa aaactcaaga tgacctgatg aaacatcaaa ccaacattag cgagctgaaa 1740  
 agaaccttct tagaaacctc aacagacact gccgtaacga atgaatggga gaagaggctt 1800  
 tccacctccc ccgtgcgact ggccgccagg caggaggatg ccccatgat cgaaccactt 1860  
 gtccctgaag agaaaatgga aaccaagacg gagtccagt gatagagacg gaaccaccg 1920  
 tgcaccacct gccgcttagc actgagaagg tgggtgcagga gaccgtgttg gtggaggagc 1980  
 ggcgtgtggt gcacgcgagt ggggatgctt cttactcggc gggagacagc ggggatgctg 2040  
 cagcacagcc cgcattcaca ggcattaaag ggaaagaggg ctctgcttga cggagggggc 2100  
 taaagaggaa ggaggggagg aggtcgctaa agctgtcctg gaacaggaag agacagccgc 2160  
 tgcttcccgt gagcgacaag aggagcagag tgcagccatc cacatttcag aaactttgga 2220  
 acaaaaacct catTTTtagt cctcaacggt gaagacggaa accatcagtt ttggcagtg 2280  
 ttcaccggga ggagtaaagc tagaaatttc cacaagaagt gccagtagtt cacaccgaaa 2340  
 ccaaaacctat cacatatgaa tcatcacagg tcgatccagg cacagatctg gagccaggcg 2400  
 tgctgatgag tgcacagacg atcacatctg aaaccaccag taccaccacc actaccaca 2460  
 tcacaaaac tgtgaaaggg ggcatttcag agacaagaat tgagaagcga atagtcatca 2520  
 cgggggatgc agacattgac catgaccagg cgctggctca ggcaattaaa gaggccaaag 2580  
 agcagcacc tgacatgtca gtgaccaaag tagtgggtcca taaagagaca gagatcacac 2640  
 cagaagatgg agaggattga ccagaggaat aacttagctt gcacatgaat gcagtcagtc 2700  
 aaaccgttag gaaaaccaga gcctatatgg agttccctct tctaaccaca ctgtacttgt 2760  
 atctgtccgt gggaaatttc agtccagaag aattgacctt gaccattaat aaagacactg 2820  
 gcagagagat cttcccataa taaagcaatc tgattcagca tctaataacc gataatgcat 2880  
 gaagcaacga taaaattaca aaagagcagc atttttaatt ttcacaaaat gtctcagttt 2940  
 tcagctatac ctgctcgttc ataaccaaca atataaaccg tgggtctcatg taacacataa 3000  
 acaattcatg cttttcatag tttattatta ttaaagtcta acaaaaattg caatttctta 3060  
 ggtaacctta tatttacaat aaatgaagat taccctcaaa tgctagaagc tgtctagggtc 3120  
 cgtccggtgt gtcagatttc ctcagattag atgtgccaat aaccaagttt attcagtaaa 3180  
 caacttgtag ttgtttcatc tggtttatta ctctcaccca taaacagtaa tgactctctg 3240  
 accctctgga aatatgtaat gcttccaatc ttgctttgtg tatctcattt aatttgttcc 3300  
 ggtaagga 3309

<210> 35

<211> 1195

<212> DNA

&lt;213&gt; Homo sapiens

<400> 35  
 ggcacgaggc gccagtcccc taacctgag gctgccgcgc ggcggtcact gcgccggggt 60  
 agtgggcccc agtggtgcgc tctctggccg ttccttacac tttgcttcag gctccagtgc 120  
 aggggcgtag tgggatatgg ccaactcggg ctgcaaggac gtcacgggtc cagatgagga 180  
 gagttttctg tactttgcct acggcagcaa cctgctgaca gagaggatcc acctccgaaa 240  
 cccctcggcg gcgttcttct gtgtggcccc cctgcaggat tttagcttg actttggcaa 300  
 ttccaaggc aaaacaagtc aaacttggca tggagggata gccaccattt ttcagagtcc 360  
 tggcgatgaa gtgtggggag tagtatggaa aatgaacaaa agcaatttaa attctctgga 420  
 tgagcaagaa ggggttaaaa gtggaatgta tgttgtaata gaagttaaag ttgcaactca 480  
 agaaggaaaa gaaataacct gtcgaagtta tctgatgaca aattacgaaa gtgctcccc 540  
 atccccacag tataaaaaga ttatttgcct gggtgcaaaa gaaaatggtt tgccgctgga 600  
 gtatcaagag aagttaaaag caatagaacc aaatgactat acaggaaagg tctcagaaga 660  
 aattgaagac atcatcaaaa agggggaaac acaaactctt tagaacataa cagaatatat 720  
 ctaagggtat tctatgtgct aatataaaat atttttaaca ctgagaaca gggatctggg 780  
 ggatctccac gtttgatccg ttttcagcag tgctctgaag gagtatctta cttgggtgat 840  
 tccttgtttt tagactataa aaagaaactg ggataggagt tagacaattt aaaaggggtg 900  
 tatgagggcc tgaaatatgt gacaaatgaa tgtgagtacc cttctgtga aactgaaag 960  
 ctattctctt gaattgatct taagtgtctc cttgctctgg taaaagatag atttgtagct 1020  
 cacttgatga tgggtgctgt gaattgtctt gctctgtctg agatttttaa aaatcagctt 1080  
 aatgagagta atctgcagac aattgataat aacattttga aaattggaaa gatggtatac 1140  
 tgtttttaga ggaataaacg tatttgtggt ttaaaaaaaa aaaaaaaaaa aaaaa 1195

&lt;210&gt; 36

&lt;211&gt; 2035

&lt;212&gt; DNA

&lt;213&gt; Homo sapiens

<400> 36  
 gaattccggg ctccggggat gaggtcgcgg ccggcgggtc ccgcgctggt gctgctgctg 60  
 ctcttcctcg gagcggccga gtcggtgcgt cgggccagc ctccgcgcg ctacacccca 120  
 gactggccga gcctggattc tcggccgctg ccggcctggt tcgacgaagc caagttcggg 180  
 gtgttcattc actggggcgt gttctcgggt ccgcctggg gcagcgagt gttctggtgg 240  
 cactggcagg gcgaggggcg gccgcagtag cagcgcttca tgcgcgacaa ctacccgccc 300  
 ggcttcagct acgccgactt cggaccgcag ttcactgcgc gcttcttcca cccggaggag 360  
 tgggccgacc tcttccaggc cgcgggcgcc aagtatgtag ttttgacgac aaagcatcac 420

gaaggcttca caaactggcc gagtcctgtg tcttgggaact ggaactccaa agacgtgggg 480  
 cctcatcgagg atttggttgg tgaattggga acagctctcc ggaagaggaa catccgctat 540  
 ggactataacc actcactctt agagtgggtc catccactct atctacttga taagaaaaat 600  
 ggcttcaaaa cacagcattt tgtcagtga aaaacaatgc cagagctgta cgaccttggt 660  
 aacagctata aacctgatct gatctgggtc gatggggagt gggaatgtcc tgatacttac 720  
 tggaactcca caaatcttct ttcattggctc tacaatgaca gccctgtcaa ggatgagggtg 780  
 gtagtaaatg accgatgggg tcagaactct tcctgtcacc atggaggata ctataactgt 840  
 gaagataaat tcaagccaca gagcttgcca gatcacaagt gggagatgtg caccagcatt 900  
 gacaagtctt cctggggcta tcgtcgtgac atggcattgt ctgatgttac agaagaatct 960  
 gaaatcattt cggaactgggt tcagacagta agtttgggag gcaactatct tctgaacatt 1020  
 ggaccaacta aagatggact gattgttccc atcttccaag aaaggcttct tgctgttggg 1080  
 aaatggctga gcatcaatgg ggaggctatc tatgcctcca aaccatggcg ggtgcaatgg 1140  
 gaaaagaaca caacatctgt atggtatacc tcaaagggat cggctgttta tgccattttt 1200  
 ctgcactggc cagaaaatgg agtcttaaac cttgaatccc ccataactac ctcaactaca 1260  
 aagataacaa tgctgggaat tcaaggagat ctgaagtggg ccacagatcc agataaagggt 1320  
 ctcttcatct ctctacccca gttgccaccc tctgctgtcc ccgcagagtt tgcttggtact 1380  
 ataaagctga caggagtga gtaatcattt gagtgcaaga agaaaggagg gctgctcact 1440  
 gttttcctgc ttcagttttt ctcttatagt accatcacta taatcaacga acttctcttc 1500  
 tccaccaga gatggctttt ccaacacatt ttaattaaag gaactgagta cattaccctg 1560  
 atgtctaaat ggaccaaaga tctgagatcc attgtgatta tatctgtatc aggtcagcag 1620  
 aagaaggaa tgagcagttg aactctgagt tcatcaattc taatatttgg aaattatcta 1680  
 caatggaatc ttcctctgt tctctgataa cctacttgct tactcaatgc ctttaagcca 1740  
 agtcaccctg ttgcctatgg gaggagggtg aaggatttgg caagctcaac cacatgctat 1800  
 ttagttagca tcagtgtga ccaacagtct ttctgcaaag ggcaggagag ctttggggga 1860  
 aaggaaaagg cttaccaggc tgctatgggt aactcttcag aaattttcag agcaatctaa 1920  
 aagcgccaaa attcgctatg ttacagtga tactattaag aaaatgaatg tgattctgct 1980  
 ctgtcttttt aagtatgatc aaataaaaaa tttgtacatc acaatcattt ctacc 2035

<210> 37

<211> 2133

<212> DNA

<213> Homo sapiens

<400> 37

cgggagagcg cgctctgcct gccgcctgcc tgcctgccac tgagggttcc cagcaccatg 60  
 agggcctgga tcttctttct cctttgcctg gccgggaggg ccttggcagc ccctcagcaa 120

gaagccctgc	ctgatgagac	agaggtggtg	gaagaaactg	tggcagaggt	gactgaggta	180
tctgtgggag	ctaactctgt	ccaggtggaa	gtaggagaat	ttgatgatgg	tgcagaggaa	240
accgaagagg	aggtggtggc	ggaaaaatccc	tgccagaacc	accactgcaa	acacggcaag	300
gtgtgcgagc	tggatgagaa	caacaccccc	atgtgctgtg	gccaggaccc	caccagctgc	360
ccagccccca	ttggcgagtt	tgagaagggtg	tgcagcaatg	acaacaagac	cttcgactct	420
tcctgccact	tctttgccac	aaagtgcacc	ctggagggca	ccaagaaggg	ccacaagctc	480
cacctggact	acatcgggcc	ttgcaaatac	atccccctt	gcctggactc	tgagctgacc	540
gaattcccc	tgcgcatgcy	ggactggctc	aagaacgtcc	tggtcaccct	gtatgagagg	600
gatgaggaca	acaaccttct	gactgagaag	cagaagctgc	gggtgaagaa	gatccatgag	660
aatgagaagc	gcctggaggc	aggagaccac	cccgtggagc	tgctggccc	ggacttcgag	720
aagaactata	acatgtacat	cttcctgtga	cactggcagt	tcggccagct	ggaccagcac	780
cccattgacg	ggtacctctc	ccacaccgag	ctggctccac	tcgctgctcc	cctcatcccc	840
atggagcatt	gcaccacccg	ctttttcgag	acctgtgacc	tggacaatga	caagtacatc	900
gccctggatg	agtgggccgg	ctgcttcggc	atcaagcaga	aggatatcga	caaggatctt	960
gtgatctaaa	tccactcctt	ccacagtacc	ggattctctc	tttaaccctc	cccttcgtgt	1020
ttcccccaat	gtttaaaatg	tttggatggt	ttgttgttct	gcctggagac	aagggtgctaa	1080
catagattta	agtgaataca	ttaacggtgc	taaaaatgaa	aattctaacc	caagacatga	1140
cattcttagc	tgtaacttaa	ctattaaggc	cttttcacac	cgcatataata	gtcccatttt	1200
tctcttgcca	ttttagctt	tgcccattgt	cttattggca	catgggtgga	cacggatctg	1260
ctgggctctg	ccttaaacac	acattgcagc	ttcaactttt	ctcttttagtg	ttctgtttga	1320
aactaatact	taccgagtca	gactttgtgt	tcatttcatt	tcagggtctt	ggctgcctgt	1380
gggcttcccc	aggtggcctg	gaggtgggca	aagggaagta	acagacacac	gatgttgtca	1440
aggatggttt	tgggactaga	ggctcagtgg	tgggagagat	ccctgcagaa	tccaccaacc	1500
agaacgtggt	ttgcctgagg	ctgtaactga	gagaaagatt	ctggggctgt	cttatgaaaa	1560
tatagacatt	ctcacataag	cccagttcat	caccatttcc	tcctttacct	ttcagtgcag	1620
tttcttttca	cattaggctg	ttggttcaaa	cttttgggag	cacggactgt	cagttctctg	1680
ggaagtggtc	agcgcctcct	gcagggttcc	tcctcctctg	tcctttggag	aaccagggtc	1740
cttctcaggg	gctctagggg	ctgccaggct	gtttcagcca	ggaaggccaa	aatcaagagt	1800
gagatgtaga	aagttgtaaa	atagaaaaag	tggagtgggt	gaatcggttg	ttctttcctc	1860
acatttggtg	gattgtcata	aggttttttag	catgttcctc	cttttcttca	ccctccccct	1920
tgttcttcta	ttaatcaaga	gaaacttcaa	agttaatggg	atggtcggat	ctcacaggct	1980
gagaactcgt	tcacctcaa	gcatttcatg	aaaaagctgc	ttcttattaa	tcatacaaac	2040
tctcaccatg	atgtgaagag	tttcacaaat	ctttcaaaat	aaaaagtaat	gacttagaaa	2100
ctgaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaa			2133

<210> 38

<211> 20

<212> DNA

<213> Homo sapiens

<400> 38

aggagggaag ggaaaacaga

20

<210> 39

<211> 20

<212> DNA

<213> Homo sapiens

<400> 39

ttaaggctca acacgaggct

20

<210> 40

<211> 20

<212> DNA

<213> Homo sapiens

<400> 40

cttgagctgt gaggtcatcg

20

<210> 41

<211> 20

<212> DNA

<213> Homo sapiens

<400> 41

tatagctcgg caccttcacc

20

<210> 42

<211> 21

<212> DNA

<213> Homo sapiens

<400> 42

ctgcctgcca ctgagggttc c

21

<210> 43

<211> 24

<212> DNA

<213> Homo sapiens

<400> 43

tccaggcaga acaacaaacc atcc

24

<210> 44

<211> 20

<212> DNA

<213> Homo sapiens

<400> 44

accaccacca ctaccacat

20

<210> 45

<211> 20

<212> DNA

<213> Homo sapiens

<400> 45

tggttttcct aacggtttgc

20

<210> 46

<211> 21

<212> DNA

<213> Homo sapiens

<400> 46

tgttggcgta caggtctttg c

21

<210> 47

<211> 19

<212> DNA

<213> Homo sapiens

<400> 47

gctacgagct gcctgacgg

19

<210> 48

<211> 24

<212> DNA

<213> Homo sapiens

<400> 48

cacattaggc tgttggttca aact

24

<210> 49

<211> 19

<212> DNA

<213> homo sapiens

<400> 49

caggatgcgc tgaccactt

19

<210> 50

<211> 19

<212> DNA

<213> Homo sapiens

<400> 50

tcctcacgcc ctgctatca

19

<210> 51

<211> 22

<212> DNA

<213> Homo sapiens

<400> 51

ttcaggatgt ccaggcatat gt

22

<210> 52

<211> 20

<212> DNA

<213> homo sapiens

<400> 52

tgtcctcatc tggaacaagg

20



<210> 53

<211> 20

<212> DNA

<213> homo sapiens

<400> 53

ggcaggagtt ctgtcctttg

20

<210> 54

<211> 20

<212> DNA

<213> homo sapiens

<400> 54

tttacaatcca gaggcacgac

20

<210> 55

<211> 19

<212> DNA

<213> homo sapiens

<400> 55

cacgatgtca gcaaacagg

19

<210> 56

<211> 20

<212> DNA

<213> homo sapiens

<400> 56

caggaaggct atggctttgg

20

<210> 57

<211> 21

<212> DNA

<213> homo sapiens

<400> 57

ccgtttcaca cctgacacat g

21

<210> 58

<211> 22

<212> DNA

<213> homo sapiens

<400> 58

gctggaccgg aagtaggttt ct

22

<210> 59

<211> 17

<212> DNA

<213> homo sapiens

<400> 59

gccgctaccg gaaatgc

17

<210> 60

<211> 20

<212> DNA

<213> homo sapiens

<400> 60

gaccagtgat atccaacaga

20

<210> 61

<211> 20

<212> DNA

<213> homo sapiens

<400> 61

gtgtgctggt aaagcttcac

20

<210> 62

<211> 22

<212> DNA

<213> homo sapiens

<400> 62

cagtaacaac caatgtgtgc ag

22

<210> 63

<211> 20

<212> DNA

<213> homo sapiens

<400> 63

tgccaaaacg acttttgaac

20

<210> 64

<211> 20

<212> DNA

<213> homo sapiens

<400> 64

gttgaccct gctaaggaaa

20

<210> 65

<211> 20

<212> DNA

<213> homo sapiens

<400> 65

cagatagttg ggcaggtct

20

<210> 66

<211> 22

<212> DNA

<213> homo sapiens

<400> 66

cactggcaaa acaatgcaga ct

22

<210> 67

<211> 22

<212> DNA

<213> Homo sapiens

<400> 67

cgaccttgac catctttgga tt

22